

INDICATIVE AND SUBJUNCTIVE CONDITIONALS

A Dissertation
Presented to the Faculty of the Graduate School
of Cornell University
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy

by
Theodore Korzukhin
January 2013

© 2013 Theodore Korzukhin
ALL RIGHTS RESERVED

INDICATIVE AND SUBJUNCTIVE CONDITIONALS

Theodore Korzukhin, Ph.D.

Cornell University 2013

Abstract. My dissertation, consisting of three independent papers, argues for a version of Stalnaker's semantics of indicative conditionals, similar to the approach taken by Heim (1992). In the first paper, I argue that competing approaches to indicatives -- the strict-conditional theories in the style of Kratzer (1986) and the approaches that rely on Adams' Thesis -- give the wrong predictions for *dominance conditionals*, sentences like 'If I win, I'll be better off than if I lose'. In the second paper, I argue that competing approaches give the wrong predictions for *singular whether-conditionals*, sentences like 'If I go, I'll go whether you like it or not'. In the third paper, I argue that the central task of a unified theory of indicative and subjunctive conditionals is to explain the presuppositions of both kinds of conditional, and defend the claim that such a unified theory is best built on the foundation of the Heimian variant of Stalnaker's indicative semantics.

BIOGRAPHICAL SKETCH

Theodore Korzukhin studied Mathematics and History at Brandeis University (1990-94). He went on to study Classics at CUNY Graduate center (1995-8), and Classics and Ancient Philosophy at Boston University (1998-2005). Moving to Cornell, he defended a dissertation in philosophy on indicative and subjunctive conditionals (2006-2012).

ACKNOWLEDGMENTS

Credit goes to Scott Sturgeon, in whose seminar at the University of Michigan I first thought that conditionals were interesting. This dissertation benefited at the early stage from the quick mind and stern advising of Brian Weatherson, who pointed me to the work of Daniel Nolan and Anthony Gillies, whose thinking, in various ways, set me on the right path. My greatest debt is to Matti Eklund, my advisor. This dissertation is mostly comments on his comments on my inchoate ideas. Harold Hodes and Nico Silins, members of my advising committee, have helped on innumerable matters, and have always kept me straight on what is reasonable and what is not. Sally McConnell-Ginet and Will Starr have helped in the closing stages of the project. My colleagues Zach Abrahams, Andrew Alwood, David van Bruwane, and David Liebesman were the primary audience of the early presentations of this work. Michelle Kosch has been my guide in many ways.

Introduction.....	1
Dominance Conditionals and the Newcomb Problem.....	15
Introduction	15
Dominance Indicatives.....	17
Kratzer's semantics.....	19
Stalnaker's semantics	24
Adams' Thesis	27
The argument in favor of Stalnaker-type semantics.....	30
The Newcomb Problem	32
Deliberationally useless indicatives	39
The semantics of indicatives and the Newcomb problem	44
Conclusion	47
Whether-Conditionals.....	56
Introduction	56
McGee's argument against modus ponens	57
Heim's semantics.....	60
Whether-conditionals	63
Taking whether-conditionals at Face Value	68
Singularity	70
The solution: Heim+	72
Why Heim?.....	78
Conclusion	81
The Unified Theory	83
Introduction	83
Kennedy pairs	86
Presuppositions.....	88
Kratzer's unified theory.....	92
Iatridou: explaining the presuppositions.....	97
My view	101
Context-dependence of subjunctives.....	105
Conclusion.....	111

Chapter 1

Introduction

In this dissertation I study three problems in the semantics of conditionals.¹ My solutions to these problems converge on a single theory, which it is appropriate to call the *neo-Stalnakerian* theory. In this introduction, I place the new approach within the history of the recent work on conditionals, sketch in a preliminary way the arguments of the three papers, and set out some background notions that help define the problems. (The discussion here will omit some important qualifications that must await until the more detailed discussion to follow; the main purpose here being to sketch the overall picture and indicate some basic notions).

The ur-theory of conditionals is the material implication theory: a conditional 'if P, then Q' is true just in case the antecedent is false or the consequent is true.² There are many reasons to be dissatisfied with the material conditional story. The standard complaint is the so-called paradoxes of material implication. So, compare two conditionals:

(1) If the sun will explode tomorrow, it will be the end of life as we know it.

and

(2) If the sun will explode tomorrow, it won't matter a bit.

1. The best philosophical introduction to the literature on conditionals is Bennett (2003), but see also Kratzer (2012) for a linguist's approach.

2. Supporters of the material implication analysis include Lewis (1976), Jackson (1987), and, more recently, Barker (1997).

But it is hard to believe that these are on par: even though it is quite probable that the sun will rise tomorrow as usual, and so quite probable that the material conditional is true (since the antecedent would be false), it does not seem probable that if the sun explodes, it won't matter a bit.³ This and similar examples have convinced most workers that the material implication theory is not viable.⁴ In response to these problems, Lewis (1976) and Jackson (1987) have produced more sophisticated versions of the material implication account. Both of these appeal to a special rule of assertibility for indicative conditionals -- the so-called Adams' Thesis. For this reason, I group these later theories with the NTV approaches (see below).

The material implication theory was an unqualifiedly unified theory of the conditional. The theories that succeeded it tended to posit important differences between two kinds of conditionals, *indicative* and *subjunctive*, with the result that, in effect, each approach to conditionals has had to offer two theories, one for each kind of conditional. The distinction between indicatives and subjunctives is a grammatical one.⁵ Indicative conditionals are conditionals in which the main verb in the antecedent and consequent is in the indicative mood: e.g. 'If it rained, it poured.' Subjunctive conditionals are conditionals in which the main verb is in the subjunctive mood: e.g. 'Had it rained, it would have poured.' In other languages (e.g. An-

3. This probabilistic way of brining out the paradoxical nature of the material implication account is due to Edgington (1995).

4. The material conditional theory also has trouble with counterfactual subjunctive conditionals (e.g. 'had I turned off the gas, there would have been no explosion'). It is a standard feature of our use of counterfactuals that at the time of utterance we know, or at least believe, that the antecedent is false. But if the material conditional story were true, we would know, or at least believe, all counterfactuals. Yet this is surely not the case.

5. For an introduction to the morphology of conditionals and useful references see Iatridou (2000). A third term frequently used, *counterfactual*, is reserved for subjunctive conditionals whose antecedents are presupposed to be false; whether all subjunctives are counterfactuals is an open question; see below for more on the notion of presupposition

cient Greek, Russian, French), the distinction is straightforward, due the the presence of a grammatically marked subjunctive. In English, the distinction is somewhat more subtle: sub-junctivity is marked with an extra layer of past tense. For example, in a counterfactual wish 'I wish I were in Hawaii' the verb of the embedded clause, 'were', is in the past tense, and yet refers to an (unreal) present circumstance -- what I wish is that something were the case in the present. The same extra marking occurs in the past: in 'Had it rained, it would have poured' the verb of the antecedent is in the pluperfect, or 'double past', tense: one of these pasts is a normal past tense, referring to a circumstance occurring before the time of utterance. But the other past is performing a different function: it is marking the mood of the verb as subjunctive. This grammatical distinction between indicatives and subjunctives turns out to be of great semantic importance: while it is possible to come up with reasonable theories for indicatives and subjunctives in isolation, it has proven remarkably difficult to produce a unified theory. And, in fact, the two kinds of conditionals seem remarkably different. For example, most theorists believe that while indicatives are context-sensitive (see below), subjunctives are not. The first and second papers of this dissertation are devoted to indicative conditionals; the last paper looks at a way in which our treatment of the two kinds of conditionals may be unified.

Two breakthroughs in the late 60's led, largely, to the abandonment of the material conditional theory, by offering alternative paradigms that dealt successfully with the problems of the material implication approach. The first breakthrough was due to Ernest Adams.⁶ On Adams' approach, the central fact about indicative conditionals is that they are assertible just

6. See Adams (1975). For a more recent expositions and defense of the Adams approach, see Edgington (1995).

when the conditional probability of the consequent given the antecedent is high (the claim known as Adams' Thesis). So Adams' approach is essentially probabilistic: indicative conditionals essentially serve to express features of the speaker's subjective probability distribution. As Lewis' triviality results showed, this approach is committed to rejecting the natural view that conditionals express propositions.⁷ For this reason, the Adams-style approaches are often called the No Truth Value theories, or *NTV* theories.

The second breakthrough was due to Robert Stalnaker.⁸ According to Stalnaker, we should understand the semantics of conditionals within a possible-worlds framework: a semantic framework that assigns truth-values to sentences not just at the actual world, but throughout the domain of all possible worlds. According to Stalnaker, an indicative conditional is true just in case the closest antecedent-world within the set of worlds consistent with what is taken for granted in the context of utterance is a consequent-world. The relative closeness of worlds is determined by a metaphysical relation of similarity between worlds (the investigation of which is subtle and ongoing).⁹

In the years following the publication of the seminal work of Adams and Stalnaker on the semantics of indicative conditionals, discussion centered on examining the relative merits of Adams-style probabilistic NTV ('no truth value') and possible-worlds approaches, and on Lewis' triviality results. It is fair to say that NTV approaches had the upper hand in that stage of the debate. But in the last twenty years the focus of discussion has largely shifted. Follow-

7. The first triviality results were given in Lewis (1976). Eells&Skyrms (1994) contains the most important of the subsequent work on triviality.

8. See Stalnaker (1975) and (1984).

9. For recent work on the similarity relation see, e.g. Kment (2006).

ing Angelika Kratzer's work on the semantics of modality and conditionals, a broad consensus emerged that a Kratzer-style strict conditional view is broadly right, and that the earlier debate has been largely superseded.¹⁰ According to the strict-conditional view, an indicative conditional 'if P, then Q' is true just in case all the worlds consistent with what is taken for granted in the context of utterance are $P \supset Q$ worlds. As a result of the shift, recent efforts have aimed, primarily, at developing and refining Kratzer's insights (for example, in the direction of dynamic semantics), rather than searching for alternatives. In fact, there has been little motivation to search for alternatives, since Kratzer's framework appeared to lack decisive challenges, and has enjoyed wide empirical success. It is fair to say that at present the field is dominated by strict-conditional and NTV views, and the Stalnaker-style semantics of indicatives has few supporters. In this dissertation I argue that there are deep reasons to reverse the trend and pursue a Stalnaker-style approach.

The distinctive feature of all the three modern theories of the indicative -- Adams', Stalnaker's, and Kratzer's -- is that each offers a *context-sensitive* semantics. What is context-sensitivity? The basic test for context-sensitivity of an expression is the following: an expression is context sensitive if there is a sentence S in which it occurs, such that on some occasions of utterance S is true, and on others it is false (assuming that all the other expressions occurring in S are not context-sensitive). That indicatives pass this test is clear from Gibbard's Sly Pete case.¹¹ Since Gibbard's case is so central to any thinking about indicative conditionals, let me present it straight away:

10. See Kratzer (1986), (2012). Gillies (2009) is a development of the view in the context of dynamic semantics.

11. See Gibbard (1981a).

Sly Pete

Sly Pete and Mr. Stone are playing poker on a Mississippi riverboat. It is now up to Pete to call or fold. My henchman Zack sees Stone's hand, which is quite good, and signals its contents to Pete. My henchman Jack sees both hands and sees that Pete's hand is rather low, so that Stone's is the winning hand. At this point the room is cleared. A few minutes later Zack slips me a note which says 'if Pete called, he won,' and Jack slips me a note which says 'if Pete called, he lost.'. . . I conclude that Pete folded. (Gibbard 1981a, p 231).

It is unproblematic that Jack's conditional is true. But, assuming that Zack knows Pete to be the sort of person who could not possibly call while knowing that his hand is the weaker one, Zack's conditional is also evidently true.¹² Assuming that 'If Pete called, he won' is false when uttered by Pete, the Sly Pete case shows that indicative conditionals pass the basic test for context-sensitivity. Cases like the Sly Pete are commonly called *Gibbard stand-offs*.

Given that an expression is context-sensitive, the next question is how to represent its meaning (in a truth-conditional semantic theory). The simplest proposal is to represent it as a function from contexts to extensions. It remains to decide how to understand what a context is. We do not need to settle this question in full generality. It is sufficient to point out that indicative conditionals seem sensitive only to what is known, or taken for granted, in the context of utterance: we can express this by saying that indicatives are sensitive to the *epistemic* context of utterance. It suffices, then, to represent the context as a set of worlds, *C*, that are compati-

12. On the Adams view, indicatives do not have truth-values, so the intuitive judgements in the Sly Pete case have to be formulated in terms of the assertibility of Jack's and Zack's conditional utterances.

ble with what is known or taken for granted. Let's call a world in which the antecedent is true an antecedent-world, and a world in which the consequent is true a consequent-world. Then, the Adams theory can be formulated as follows: an indicative is assertible iff most antecedent-worlds in C are consequent-worlds.¹³ The Stalnaker theory can be formulated as follows: an indicative is true iff the closest antecedent-world in C is a consequent-world. And Kratzer's theory can be formulated as follows: an indicative is true just in case all the antecedent-worlds in C are consequent-worlds. The last question that deserves consideration is whether to understand C as the set of worlds that are compatible with what is known, or as the set of worlds compatible with what is taken for granted in the context of utterance. This is a subtle question, and I will not take sides on it in this work (for discussion, see Nolan (2003)).

In the first and second paper of my dissertation I argue against Kratzer-style and NTV approaches, and defend and develop a variant of Stalnaker's semantics for indicatives first proposed by Irene Heim (1992). I also explore the interconnections between views on the semantics of conditionals and work on the Newcomb problem and the foundations of decision theory. In the third paper, I argue for a new desideratum for a unified theory of indicative and subjunctive conditionals, and sketch a unified theory that meets it, building on the account of indicatives defended in the first two papers. A brief sketch of the main arguments follows.

There are, then, three very different competitor theories: Adams', Stalnaker's, and Kratzer's. The history of the last 40 years has shown that it is very hard to produce decisive ar-

13. It is natural to think that the subjective probability function that figures in Adams' thesis assigns non-zero probability only to worlds consistent with what the subject knows, or takes for granted, in a context -- thus the present formulation of Adams' theory is equivalent to that I gave earlier, in terms of Adams' thesis.

guments in favor of one or the other of the three competitor theories. But it is often the case that a given linguistic construction does not show all of its semantic features when considered only in un-embedded sentential contexts.¹⁴ And indeed embedded conditionals display interestingly complex behavior (see, e.g. the discussion of McGee's counter-example to *modus ponens* in the second paper). In the first paper, 'Dominance Conditionals,' I show that the consideration of a particular kind of embedded context allows for a strong argument in favor of Stalnaker's semantics, and against its competitors.

In 'Dominance Conditionals' I show that both Kratzer-style theories and NTV theories make the wrong predictions on certain *dominance conditionals*, sentences like 'If I win I'll be better off than if I lose.' Consider the following scenario:

Tea-Party

Two cups stand on the table, cup A and cup B. We do not know how much liquid is in each cup, but we do know that cup A has more liquid than cup B.

Plausibly, one ought to accept that the following (indicative) dominance conditional is true in *Tea-Party*:

- (3) If I drink the contents of cup A, I will ingest more than if I drink the contents of cup B.

I show that, on plausible assumptions about the semantics of comparatives, a Stalnaker-style view predicts that (39) is true in *Tea-Party*, that Kratzer-style views predict that (39) is false, and that NTV theories predict that (39) is unassertable. We have a strong intuition that (39) is

14. The reason is that the contrast between asserted and unasserted clauses can be very informative: when we consider an asserted clause, our intuitions about its content are mixed with whatever the theory of assertibility tells us. This interference effect is absent when the relevant clause is embedded. Thus, the study of embeddings is in general a purer science.

true in *Tea-Party*; so this is an argument for preferring a Stalnaker-like semantics to strict-conditional and probabilistic approaches.

Dominance conditionals are also crucial to decision theory. One way to approach the Newcomb problem is through consideration of the natural-language reasoning that leads to the one-boxing and the two-boxing answers to the problem. But although David Lewis (1986) famously suggested that one-boxers are guided by indicative conditionals, and that two-boxers are guided by subjunctives, the exact nature of that reasoning has not been investigated. *How* do conditionals guide us to one or the other answer? I argue first that it is quite plain that we rightly use indicatives in deliberation. So, Lewis offers the wrong dichotomy: the question is not which kind of conditionals to appeal to (it should be indicatives), but rather which theory of indicatives underlies our reasoning to the one- or the two-boxing answer. Second, I argue that this reasoning appeals to a kind of dominance norm (*DN*), to the effect that if the relevant dominance indicative conditional is known to be true, then the course of action it 'recommends' is the rational course of action. Since

(4) If I choose two boxes, I will get more than if I choose one box.

is known to be true in the Newcomb scenario, given Stalnaker's semantics, it follows that Stalnaker's semantics together with *DN* provides an argument for two-boxing. But I argue that no *DN*-like principle plausibly connects an Adams semantics for indicatives with the one-boxing answer: it appears, *pace* Lewis, that conditionals do not guide us to one-boxing at all.

In the second paper, 'Whether-conditionals', I argue that a suitably enriched Heimian semantics, itself a version of Stalnaker's approach, alone can explain the puzzling phenomenon

of *singular whether-conditionals*. Start with the following scenario:

Pass/Fail

John is about to take an exam, but he does not like his chances of passing. He is considering praying for divine help.

Now consider the following assertion (a singular whether-conditional), made by a friend of John's:

(5) If John passes, he'll pass whether he prays or not.

Intuitively, if (5) is true in *Pass/Fail*, then we may infer that the world of the *Pass/Fail* scenario does not contain a *vengeful* god: a god who punishes those who do not pray. Likewise, we may infer that there is no *beneficent* god: a god who rewards those who pray. The question is: how to explain these readings? These readings are highly surprising, since (5) is predicted to be trivially true and so uninformative by virtually every theory of indicative conditionals, because virtually every theory accepts the principle of *Import-Export* (--the equivalence of 'If P, then if Q, then R' and 'If P and Q, then R').

I argue that the only view that explains these readings is a combination of the semantics proposed by Heim (1992) with two further interesting assumptions: one, that indicatives carry an open-consequent presupposition, and, two, that, within a dynamic account of presupposition projection, we have recourse to an up-accommodation mechanism (a mechanism that accommodates presuppositions by expanding the epistemic context, rather than by restricting it). The end result is a new argument for Heim's version of Stalnaker's semantics, augmented by a new mechanism of presupposition accommodation.

The second paper appeals to the notion of presupposition, and this notion occupies center-stage in the third paper. What are presuppositions? The most basic test for presuppositional meaning is the negation test. So, from

(6) It's the knave that stole the tarts.

it follows that

(7) Somebody stole the tarts.

but, surprisingly, (10) also seems to follow from

(8) It isn't the knave that stole the tarts.

But note that if our logic is classical and 'follows' is a consequence relation connecting the semantic contents expressed by (15)-(63), it follows that (10) is logical truth. This is an implausible conclusion. Different reactions to this observation are possible. According to the Frege-Strawson response, such examples show that our logic is not classical, but instead three-valued (see Strawson 1950, and Beaver 2001). It is natural to think of the third value as 'undefined': when a sentence has this value, it fails to express a proposition. So, when (10) is false, neither (15) nor (63) have a classical truth-value. It is thus natural to say that (10) is presupposed by both (15) and (63). According to this response, presuppositions are definedness conditions: conditions under which a sentence has a classical truth-value.

Another possible reaction (see Stalnaker 1998) is to deny that presuppositions are semantic entailments: conditions on the sentence's having a classical truth-value (but the story is compatible with some presuppositions being semantic). Rather, presupposition is for Stalnaker a pragmatic phenomenon, a matter of what speakers, literally, *presuppose*, or take for

granted in a conversation. On this diagnosis, (10) is not a logical consequence of (15): they are tied by a weaker relation (we may pragmatically infer (10) from (15)).

However one diagnoses the problem presented by (15)-(63), it is agreed on all sides that such examples present to us a new phenomenon: the phenomenon of presupposition. The negation test is widely taken to be the paradigmatic test for presuppositionhood (so (10) passes this test: (10) is presupposed by both (15) and (63)).¹⁵

It is widely agreed that indicative conditionals presuppose that the antecedent is not known to be false -- that it is epistemically possible. Likewise, it is widely agreed that nearly all subjunctives presuppose that their antecedent is known to be false.¹⁶ Given that, *prima facie*, we would like to have a unified semantics for both indicative and subjunctive conditionals, this divergence in the presuppositions of the two constructions calls for an explanation. This is the starting-point for the final paper.

In 'A Unified Theory of Conditionals,' I start with the proposal that one central task for a unified theory of the conditional is to explain the pattern of presuppositions of indicative and subjunctive conditionals. The key phenomenon is the counterfactuality presupposition of subjunctive conditionals. So,

15. Although the negation test is paradigmatic, it is not always easily applicable. As a result, the notion of presupposition is usually defined in more complicated ways (see Beaver 2001 for discussion). Indeed, Beaver writes: 'So what is the defining characteristic of the recent linguistic study of presupposition? We will see that a large class of lexical items and grammatical constructions, including those identified as presuppositional by philosophers such as Frege and Strawson, produce distinctive patterns of inference. It is difficult to find any common strand to current analyses of presupposition, save that they all concern (various parts of) this class. (Beaver 2001, p. 10).

16. Examples of the sort first introduced by Anderson (1950) are thought to show that this generalization is not strict: some subjunctives are assertible (and true) even though the truth of the antecedent is compatible with what is known.

(9) Had Oswald not killed Kennedy, someone else would have.

presupposes that Oswald killed Kennedy. The semantics of subjunctive conditionals by itself cannot explain why they carry the counterfactuality presupposition. Another hypothesis is that the pattern of conditional presuppositions can be explained by appeal to the subjunctive and indicative morphology, along the lines suggested by Iatridou (2000). I show that this approach succeeds only on the condition of being combined with an appropriately unified semantics of the conditional. In other words, Iatridou's proposal imposes a constraint on the unified semantics of the conditional. This constraint is not met by any viable existing theory. By considering some variants on Stalnaker's approach I suggest that the constraint can be met if we posit a context-sensitive modal base in the semantics of the conditional.

However, the current wisdom is that subjunctive semantics does not need to appeal to such a context-sensitive modal base. The current view, instead, is that subjunctives always report on the same unrestricted metaphysical domain of possible worlds. I show that the current wisdom is mistaken by producing a counterfactual Gibbard stand-off (see above) -- a proof that subjunctives, just like indicatives, are sensitive to what is known in the context of utterance.

The argument from the need to explain conditional presuppositions, and the argument for context-sensitivity of subjunctives thus reinforce each other and converge on a single unified semantics, which 1) explains the presuppositional pattern of indicatives and subjunctives; 2) predicts subjunctive stand-offs; and 3) is built on the foundation of the Lewis-Stalnaker similarity relation.

The dissertation as a whole, then, progresses bit by bit toward a new unified theory of the conditional. In the first paper, I argue that Stalnaker-style theories of indicatives are preferable to other alternatives. In the second paper, I show that of the Stalnaker-style theories, a version of Heim's approach is the most promising. Finally, the third paper argues that one can build a unified theory of both indicatives and subjunctives on the basis of my preferred version of Heim's approach.

Chapter 2

Dominance Conditionals and the Newcomb Problem

1 Introduction

In this paper I aim to make progress on two fronts, the semantics of indicative conditionals and the Newcomb problem, by investigating an interesting type of sentence, the *indicative dominance conditional*. There are two main claims.

The first claim is that evidence drawn from indicative dominance conditionals favors a Stalnaker-type¹⁷ semantics for indicative conditionals over its main rivals -- the strict-conditional semantics due to Angelika Kratzer and the probabilistic theories relying on Adams' Thesis (*Adams theories*, for short). I will argue for this by investigating *indicative dominance conditionals*, sentences like the following:

- (10) If I drink the (entire) contents of cup A, I will ingest more than if I drink the (entire) contents of cup B.

Intuitively, (10) is true in any situation in which we know that cup A has more tea than cup B.

17. I lay out a version of Stalnaker's original (1975) semantics for indicatives in §4. But there are several views that agree with Stalnaker's semantics in simple cases, yet diverge over more complicated conditional constructions (e.g. McGee (1985), and Heim (1992)). I will call the family of theories that agrees with Stalnaker's truth-conditions for what I call *straight* indicatives *Stalnaker-type theories*, where a straight indicative has the form 'if P, then Q', where P and Q do not involve any modal material (in particular, no nested conditionals). The argument I present in §2-6 is not an argument for Stalnaker's particular approach, but rather for a whole family of theories -- the Stalnaker-type theories. The question which Stalnaker-type theory is preferable requires separate treatment.

Of particular interest are *variably dominant* scenarios: cases where we know that cup A has more tea than cup B, but there is no amount X such that we know, of X, that the amount in cup A is greater than X and the amount in cup B less. I will argue that Kratzer's semantics predicts that (10) is false in some variably dominant scenarios, and that Adams theories predict that (10) is unassertible in some such scenarios. But Stalnaker's semantics predicts that (10) is true in all such cases. So, I claim, the predictions of Stalnaker-type semantics fit better with our intuitions. Since Stalnaker-type semantics is the only theory of indicatives that makes the right predictions, dominance indicatives provide an argument in favor of Stalnaker-type semantics, and against its rivals (that's the first main claim).

It has long been recognized that conditionals play a central role in motivating both the one-boxing and the two-boxing answers to the Newcomb problem. But the question *how* conditionals guide us to these answers has not been posed, or answered. The question is of interest independently of whether one is convinced that one-boxing, or two-boxing, is the right answer. My second main claim is that the *how* question is answered by a principle I will call the *Dominance Norm (DN)*, connecting the truth of dominance indicatives with rational courses of action. Given *DN*, our answer to the Newcomb problem can be tightly correlated with our views on the semantics of indicative conditionals, specifically, with our predictions for the Newcomb dominance conditionals:

(11) If I choose one box, I will win more than if I choose two boxes.

(12) If I choose two boxes, I will win more than if I choose one box.

(call these the *one-boxing* and the *two-boxing dominance conditionals*, respectively)

After defending DN against a potential objection stemming from Keith DeRose (2010), I will look at what Adams theories and Stalnaker's semantics tell us about the Newcomb problem in the light of DN. In particular, I argue that Stalnaker's semantics for indicatives predicts that (12) is known to be true, and so commits one to two-boxing *via* DN. Combined with the argument for a Stalnaker-type semantics, this yields a new argument for two-boxing in the Newcomb scenario.¹⁸ But DN does not yield a sound argument for one-boxing on the supposition that the Adams semantics is right, contrary to a widely cited suggestion by Lewis (1981a). If DN is indeed the bridging principle that connects conditionals and rational deliberation, one-boxers cannot appeal to conditionals to help their case.

2 Dominance Indicatives

I am going to argue that, of the dominant theories of indicative conditionals, only Stalnaker-type semantics gives plausible verdicts for the type of dominance indicatives I will call *variably dominant*. I will restrict my attention to three prominent views on the semantics of indicatives: Stalnaker (1975), Kratzer (1986), and Adams theories -- the theories based on Adams' Thesis, like Edgington (1995).¹⁹ I propose to investigate the predictions of the three theories for *in-*

18. Does the two-boxing answer need a new argument in its favor? The recent writers on the Newcomb problem tend to agree that two-boxing is the right answer (e.g. Egan (2007), Wedgwood (2011), DeRose (2010), Joyce (2007)). But the recent unanimity does not go beyond acknowledging the intuitive force of Lewis' influential discussion of the problem (Lewis (1981a, b)). For those already in the two-boxing camp, the main point of my argument in §7-9 is to clarify the connections between semantics and decision theory, and pin-point the sources of our two-boxing intuitions.

19. Although I explicitly consider only Kratzer's work below, I will argue that the problematic feature of her account is its strict conditionality. Thus there is reason to believe that my argument extends to any strict-conditional theory of indicatives, such as the dynamic view of Gillies (2009). For reasons of space, I will not attempt to generalize my argument.

dicative dominance conditionals, sentences like the following:

- (10) If I drink the (entire) contents of cup A, I will ingest more than if I drink the (entire) contents of cup B.

When is (10) true? As it turns out, the three theories under investigation deliver interestingly different verdicts. I will proceed by comparing the predictions of the three theories for (10) in a few hypothetical scenarios. The main argument will revolve around the following scenario:

Tea-party 1 (T1)

Two cups of tea are on the table -- cup A and cup B. We know that either cup A contains 2ml and cup B 1ml of tea, or cup A contains 4ml and cup B 3ml of tea.

The distinguishing feature of T1 is that, in T1, we know that cup A has more tea than cup B, but there is no amount X, such that we know, of X, that cup A has more tea than X, and cup B less. Let us call such a case a case where the amount in cup A *variably dominates* the amount in cup B. In T1, one amount dominates another in the *epistemic context of utterance* -- that is, throughout the set of worlds consistent with what is known, or taken for granted, in the context of utterance.²⁰

The basic intuition that I am going to appeal to is that (10) is true in T1. Of course, the intuition is clearly stronger than that, since it is clear that (10) should also be true in *all* the cases where the amount in cup A dominates the amount in cup B. Now let's turn to our three theories.

20. Nolan (2003) argues that we should understand the epistemic context as the set of worlds consistent with what is *known*. I am convinced by his argument, but the choice between 'known' and 'taken for granted' will not matter below.

3 Kratzer's semantics

Let's start with Kratzer. Kratzer gives the following truth-conditions for what I will call *straight indicatives*, that is, non-nested indicatives without an overt modal in the consequent:

Kratzer's semantics for straight indicatives

‘If P, then Q’ is true just in case all P-worlds in C are Q-worlds.²¹

(Where C is the epistemic context of utterance, or the *modal base*.)

Since dominance conditionals are comparative constructions, we need a semantics for comparatives to derive Kratzer's predictions for T1. Unfortunately, the semantics of comparatives is a matter of some dispute.²² For the sake of simplicity of exposition, I will adopt Schwarzschild's (2008) proposal, since its presentation does not need the apparatus of formal semantics. In the *Appendix*, I show that the relevant predictions are the same on the more standard degree-based and the alternative interval-based approach to the semantics of comparatives.

On Schwarzschild's *A-not-A* account, the meaning of

(13) A is more expensive than B.

21. The lexical entry for the epistemic ‘if’ that gives rise to these truth-conditions is something like $\lambda P_{\langle s, t \rangle}. \lambda Q_{\langle s, t \rangle}. \forall w \in C (P(w) \supset Q(w))$ -- a generalized quantifier over worlds in C. Kratzer's general theory of the indicative is more complicated, but we shall not need it, since our focus is on epistemic indicatives only.

22. For some recent contributions to the debate, see Aloni&Roelofsen (2011), Beck (2010), Schwarzschild (2008), Rooij (2008).

is given by

- (14) There is some expense threshold θ such that A meets or exceeds θ , but B does not meet or exceed θ .

The following two examples from Schwarzschild 2008 illustrate an interesting flexibility of the threshold approach when it comes to dealing with quantifiers in the than-clause:

- (15) The balloon is higher today than it has been on any other day.
 (16) The balloon is higher today than it has been on at least one other day.

According to the A-not-A analysis, (15) and (63) both say that the balloon met or exceeded a threshold θ today. But, further, (15) says that

- (17) The balloon did **not** meet or exceed θ on any other day.

while (63) says that

- (18) There was at least one other day on which the balloon did **not** meet or exceed θ .

Schwarzschild's proposal is that in order to provide adequate truth-conditions for comparatives with quantifiers in the than-clause, the A-not-A theory needs to allow the quantifiers to be sometimes interpreted inside (as in (64)) and sometimes outside the negation (as in (18)). The position outside the negation is the normal case. But negative polarity items and some modals are interpreted inside the negation. Here is an example with a modal:

- (19) The balloon is higher than it is allowed to be.

(65) says that the balloon meets or exceeds a threshold that it is **not** allowed to meet or exceed.

Now, consider the dominance conditional

- (10) If I drink the (entire) contents of cup A, I will ingest more than if I drink the (entire) contents of cup B.

The A-not-A approach suggests two possible logical forms, depending on whether the negation out-scopes the conditional in the than-clause or not. On one reading, (10) asserts that

- (20) There is a threshold θ such that if I drink the contents of cup A, I will ingest an amount that meets or exceeds θ , and it is **not** the case that if I drink the contents cup B, I will ingest an amount that meets or exceeds θ .

On the second reading, (10) asserts that

- (21) There is a threshold θ such that if I drink the contents of cup A, I will ingest an amount that meets or exceeds θ , and if I drink from cup B, I will **not** ingest an amount that meets or exceeds θ .

(20) and (21) are independent of any particular semantics of indicatives. What are the predictions when (20) and (21) are combined with Kratzer's theory? Combined with Kratzer's semantics, (20) gives:

- (22) There is a threshold θ such that for all the worlds in C in which I drink from cup A, I ingest an amount that meets or exceeds θ , and it is not the case that for all worlds in C in which I drink from cup B, I ingest an amount that meets or exceeds θ .

(22) is true just in case the *minimum* amount I drink in the worlds in which I drink from cup A is greater than the *minimum* amount I drink in the worlds in which I drink from cup B.

On the other hand, (21), combined with Kratzer's theory, gives:

(23) There is a threshold θ such that for all the worlds in C in which I drink from cup A, I ingest an amount that meets or exceeds θ , and for all worlds in C in which I drink from cup B, I do not ingest an amount that meets or exceeds θ .

(23) is true just in case the *minimum* amount I drink in the worlds in which I drink from cup A is greater than the *maximum* amount I drink in the worlds in which I drink from cup B.

But now consider the predictions of (22) and (23). First, it is clear that (22) does not work: that one minimum is greater than the other is no guarantee that if I drink from cup A I will ingest more than if I drink from cup B. The following scenario illustrates the problem:

Tea-party 2

Two cups of tea are on the table -- cup A, and cup B. We know that cup A contains at least 10ml of tea, and cup B at least 1ml of tea, and nothing else of relevance (in particular we do not know that cup A contains more tea than cup B).

Clearly, (10) may well be false in T2. Yet (22) predicts that (10) is true. This is the wrong result.

What about (23)? (23) predicts that (10) will be false in T2, which is perhaps the right result.²³ But (23) makes the wrong prediction for T1:

Tea-party 1

Two cups of tea are on the table -- cup A and cup B. We know that either cup A contains 2ml and cup B 1ml of tea, or cup A contains 4ml and cup B 3ml of tea.

Here the A-minimum is 2ml, and the B-maximum is 3ml, so (10) comes out false according to

23. It is *perhaps* the right result because it may be that the description of the T2 scenario underspecifies the situation, so that (10) is true in some situations conforming to the description in T2, and false in others.

(23). Yet it is clearly true that, in T1, if I drink from cup A I will ingest more than if I drink from cup B. So (23), too, gives the wrong predictions.

One might think that (20) and (21) do not exhaust the possible logical forms of (10) on Schwarzschild's view. (20) and (21) both interpret the comparative as out-scoping both conditionals, so that the result is a comparison of the amounts I ingest if I drink from cup A with the amounts I ingest if I drink from cup B. But one might think that perhaps the right logical form for (10) has the comparative inside the consequent of the first conditional, so that the logical form looks as follows:

(24) If I drink from cup A [I will ingest more than if I drink from cup B]

So there are two more options to consider. The A-not-A approach gives two possible interpretations. One is:

(25) If I drink the contents of cup A, [then there is a threshold θ such that I will ingest an amount that meets or exceeds θ , and if I ingest from cup B, I will not ingest an amount that meets or exceeds θ .]

and the other is

(26) If I drink the contents of cup A, [then there is a threshold θ such that I will ingest an amount that meets or exceeds θ , and it is not the case that if I ingest from cup B, I will ingest an amount that meets or exceeds θ .]

But (60) is vacuously true when combined with Kratzer's semantics. The vacuity is independent of the contribution of the comparative, and is the same as the vacuity of

(27) If I drink from cup A, then if I drink from cup B I will ingest some tea.

Since we are supposing that the drinking is a single event, the first and second antecedent of (27) are inconsistent: there are no epistemically possible worlds in which I drink from both cups. The nested conditional 'if I drink from cup B I will ingest some tea' is evaluated against a modal base that contains only the worlds in which I drink from cup A, and so (27) comes out vacuously true. The same goes for (60): combined with Kratzer's semantics, it is vacuously true. When combined with Kratzer's semantics (61) is, for the same reason, vacuously false. So the alternative logical forms (60) and (61) do not help to resolve the problem.

The argument above shows that the combination of Kratzer's semantics with Schwarzschild's account of comparatives makes the wrong predictions. But could it be that it is the semantics of comparatives that is in need of revision? Part of the answer is contained in the *Appendix*, where I consider two other prominent theories of the comparative, and show that they yield the same conclusions I have reached above. Furthermore, in §6 I will offer a diagnosis of the problem that suggests that it is the strict-conditionality of Kratzer's view that is the culprit. Kratzer's semantics gives the wrong predictions for variably dominant dominance conditionals.

4 Stalnaker's semantics

Now, let us turn to Stalnaker-type theories. Stalnaker's (1975) semantics for indicatives relies on the similarity metric familiar from the standard Stalnaker-Lewis semantics for counterfac-

tuals.²⁴ The simplest formulation goes as follows:

*Stalnaker's semantics*²⁵

An indicative 'if P then Q' is true at w iff the closest P-world to w in C is a Q-world.

(here C is the epistemic context of utterance: the set of worlds consistent with what is taken for granted, or known, in the context of utterance).

It should not be surprising that Stalnaker's semantics predicts the truth of (10) in T1. First, recall that the A-no-A theory gives two possible logical forms for (10):

- (20) There is a threshold θ such that if I drink the contents of cup A, I will ingest an amount that meets or exceeds θ , and it is **not** the case that if I drink from cup B, I will ingest an amount that meets or exceeds θ .

and

- (21) There is a threshold θ such that if I drink the contents of cup A, I will ingest an amount that meets or exceeds θ , and if I drink from cup B, I will **not** ingest an amount that meets or exceeds θ .

Combined with Stalnaker's semantics, the two logical forms yield:

- (28) There is a threshold θ such that the closest world w in which I drink the contents of cup

24. I incorporate the uniqueness assumption, the assumption that for all worlds w, there always is a unique closest P-world to w, into the semantics. Although I think that there are good reasons for accepting the uniqueness assumption, nothing in what follows hinges on it.

25. The view defended in Stalnaker 1975 does not include the explicit restriction on the context incorporated in semantics in the text. The main motivation for formulating things as I do is to avoid complicated questions about the presuppositions of indicatives that Stalnaker's actual view depends on. In any case, our interest is in Stalnaker-type theories, and the variant I offer is just the simplest kind.

A is such that, in w , I drink an amount that meets or exceeds θ , and it is not the case that the closest world w' in which I drink from cup B is such that, in w' , I drink an amount that meets or exceeds θ .

and

- (29) There is a threshold θ such that the closest world w in which I drink the contents of cup A is such that, in w , I drink an amount that meets or exceeds θ , and such that the closest world w' in which I drink from cup B is such that, in w' , I do not drink an amount that meets or exceeds θ .

But both (28) and (29) yield the same result: both are true in T1.²⁶ So, suppose that we are evaluating (10) at a world w in accordance with (28). Suppose that w is a 2ml-1ml world in which I drink from cup A. Then there is a threshold, 2ml for example, such that the closest A-world (which, by hypothesis, is w itself) is such that in it I drink an amount that meets the threshold. What about the closest B-world? Here one has to appeal to the central feature of the Lewis-Stalnaker similarity metric -- that it tracks (more or less) causal dependence and independence. In particular, it is plausible that the closest world to w in which I drink from cup B is a 2ml-1ml world, just like w itself. With that assumption, it follows that in the closest B-world to w I drink an amount that does not meet 2ml -- our chosen threshold. So (10) is true in w on Stalnaker's semantics. But this reasoning generalizes to all the worlds in C. So (10) is true throughout C. Stalnaker's semantics gives the right predictions for variably dominant

26. Recall that I am assuming, for simplicity, that there are no ties for closest world. If ties are allowed, (28) and (29) may diverge.

scenarios.²⁷

5 Adams' Thesis

Adams' Thesis is the claim that:

Adams' Thesis

An indicative 'if P, then Q' is assertible iff $p(Q|P)$ is high.

There are two kinds of theory that explicitly subscribe to Adams' Thesis -- the material conditional accounts of Lewis (1976) and Jackson (1987), and the so-called *No-Truth-Value* theories of Edgington (1995) and Bennett (2003). But in what follows, all that matters is Adams' Thesis itself, so I will just speak of all the theories that subscribe to Adams' Thesis, collectively, as *Adams theories*. Some Adams theorists deny that indicative conditionals have truth-values, preferring to speak about their assertibility in context, and this creates severe difficulties for deriving the predictions of such an approach for dominance indicatives, because one cannot appeal to the standard compositional framework of truth-conditional semantics, as I have done above. So to derive the predictions of Adams theories one must use a more indirect approach.

What kind of constraints are there on what the Adams theorist can say about dominance indicatives? First, it seems that the assertibility of (10) can depend only on conditional probabilities $p(\text{I will ingest } d\text{-much} \mid \text{I drink the contents of cup } A/B)$ for all d . The intuition behind this assumption is simple: whatever the exact Adams semantics of the dominance conditional

27. For T2, the prediction is that (10) may be true or false, depending on the amounts of tea in A and B -- that, too, is intuitively the right result.

is, it ought to be compositional, and a dominance indicative is somehow composed of two conditional constructions and a comparative. According to Adams theorists, the contribution of the conditional is the conditional probability, hence the assumption that the assertibility of (10) can depend only on conditional probabilities $p(\text{I will ingest } d\text{-much} \mid \text{I drink the contents of cup } A/B)$ for all d . Let's call this assumption the *minimal assumption of compositionality*.

Further, it is reasonable to assume that Adams dominance conditionals are *extensional*, in the sense that their assertibility-value depends only on the conditional probabilities of the subject's *current* probability distribution, and not on any other probability distributions.²⁸

Of course, even with the extensionality and compositionality assumptions accepted, we still need an answer to the question *how* the conditional probabilities just mentioned determine the assertibility of (10). In the absence of a compositional semantics, one could try to guess the assertibility-conditions that an Adams theorist might wish to assign to dominance indicatives. I do not have any good candidates to offer.²⁹ But it can be shown that *any* Adams semantics for dominance conditionals will have problems with variably dominant cases, given the extensionality and compositionality assumptions.

Together, the two assumptions amount to the assumption that the assertibility of (10) de-

28. It is reasonable to think that (10) is equivalent to

The amount I will ingest if I drink the contents of cup A is greater than the amount I will ingest if I drink the contents of cup B.

If this is right, the extensionality assumption looks very natural, for it is clear that this comparative depends only on the current probability distribution.

29. Here is one *prima facie* reasonable guess: The dominance conditional 'If I drink the contents of cup A, I will ingest more than if I drink the contents of cup B' is assertible just in case the maximum d for which 'If I drink the contents of cup A, I will ingest at least d -much' is assertible, is greater than the maximum d for which 'If I drink the contents of cup B, I will ingest at least d -much' is assertible. (in terms of probabilities: just in case $\max \{d; p(\text{I will ingest at least } d\text{-much} \mid \text{I will drink the contents of cup A}) \text{ is high}\} > \max \{d; p(\text{I will ingest at least } d\text{-much} \mid \text{I will drink the contents of cup B}) \text{ is high}\}$)

depends only on the conditional probabilities $p(\text{I will ingest } d\text{-much} \mid \text{I drink the contents of cup } A/B)$ of the current probability distribution $p()$. Given this assumption, the argument is straightforward. Compare T1:

Tea-party 1

Two cups of tea are on the table -- cup A and cup B. We know that either cup A contains 2ml and cup B 1ml of tea, or cup A contains 4ml and cup B 3ml of tea.

with

Tea-party 3

Two cups of tea are on the table -- cup A and cup B. We know that either cup A contains 2ml and cup B 3ml of tea, or cup A contains 4ml and cup B 1ml of tea.

Let us stipulate further that in both T1 and T3 the relevant possibilities are equiprobable: so that

$$\begin{aligned}
 & p_{T1}(\text{I drink from cup A and cup A has 2ml and cup B has 1ml}) \\
 &= p_{T1}(\text{I drink from cup B and cup A has 2ml and cup B has 1ml}) \\
 &= p_{T1}(\text{I drink from cup A and cup A has 4ml and cup B has 3ml}) \\
 &= p_{T1}(\text{I drink from cup B and cup A has 4ml and cup B has 3ml}) = .25
 \end{aligned}$$

in T1, and similarly in T2:

$$\begin{aligned}
 & p_{T3}(\text{I drink from cup A and cup A has 2ml and cup B has 3ml}) \\
 &= p_{T3}(\text{I drink from cup B and cup A has 2ml and cup B has 3ml}) \\
 &= p_{T3}(\text{I drink from cup A and cup A has 4ml and cup B has 1ml}) \\
 &= p_{T3}(\text{I drink from cup B and cup A has 4ml and cup B has 1ml}) = .25
 \end{aligned}$$

T1 is variably dominant, while T3 is not. It is clear that (10) is true (assertible) in T1, and false (unassertible) in T3. But in T1 and T3 the relevant conditional probabilities are identical: $p_{T1}(\text{I will ingest d-much} \mid \text{I drink the contents of cup A/B}) = p_{T3}(\text{I will ingest d-much} \mid \text{I drink the contents of cup A/B})$. So if (10) is not Adams-assertible in T3, it is not Adams-assertible in T1, either.

To sum up: although we do not have a direct route to a compositional Adams semantics for dominance conditionals, the argument appealing to T1 and T3 shows that it is bound to give the wrong predictions, no matter what it is, so long as it respects the two assumptions of minimal compositionality and extensionality made above.

6 The argument in favor of Stalnaker-type semantics

Our basic result, then, is this: Kratzer's semantics and Adams theories give the wrong verdict on (10) in T1 (or T2 or T3), while Stalnaker predicts that it is true (and assertible) in T1, false in T3, and true or false in T2, depending on the actual amounts in A and B. Since (10) is, plausibly, true (and assertible) in T1, false in T3, and true or false depending on the amounts in A and B in T2, we have an argument in favor of Stalnaker's semantics.

Intuitively, the reasoning in §3-5 is a dramatization of the following basic fact. Both Kratzer's semantics and Adams theories are what one might call *global*: the truth/assertibility of an indicative conditional is determined entirely by the epistemic context of utterance, and is insensitive to which world in the epistemic context is the world of evaluation. Stalnaker's

semantics, on the other hand, is what one might call *local*: it is sensitive both to the epistemic context, and to the identity of the world of evaluation. Because they are global, Kratzer and Adams dominance indicatives report, in T1, on the relations between the entire set of amounts I ingest in the worlds in which I drink from cup A, and the entire set of amounts I ingest in the worlds in which I drink from cup B. But, in variably dominant scenarios like T1, no such relation -- neither the relation between the minima of both sets, nor between minima and maxima -- is informative enough; no such relation tracks the truth of dominance conditionals in variably dominant scenarios. The argument against Adams theories above demonstrates this particularly clearly, by appealing to a pair of scenarios (T1 and T3) which are indistinguishable from the point of view of any Adams theory of dominance indicatives. The argument against Kratzer's view above is also very similar, although the details about the semantics of comparatives may obscure this: the argument works, essentially, by showing that the Kratzer semantics *does not pass on enough information* to the semantics of comparatives, and, as a result, the Kratzer dominance conditionals are not sensitive enough to distinguish between the peculiar variably dominant cases, in which (10) is true, and the ordinary non-dominant ones, in which (10) is false.

Local theories, by contrast, *can* give the comparative semantics enough information. So, for Stalnaker, the truth of (10), evaluated in a given world w in the epistemic context, depends on a comparison between w and another world (the closest world in which I drink from the other cup). And so (10) is *assertible*, or *known*, if (10) is true in most, or in all the worlds in the epistemic context. Thus, while global theories track the relations between sets of degrees (=the

amounts I ingest in A-worlds and B-worlds), local theories like Stalnaker's track the sets of pairwise relations between worlds. If this diagnosis is right, it seems plausible that no future semantics of comparatives could rescue global theories like Kratzer's account (or Adams theories).

7 The Newcomb Problem

Recall the Newcomb problem. Two boxes stand in front of John. One is transparent, and contains \$1000. The other is opaque, and contains either nothing or \$1,000,000. John is offered a choice -- pick the opaque box ('pick one box'), or pick both boxes. John is to receive the money contained in the boxes he chooses. John is also told that a very reliable (say, .99) computer has already set the amount in the opaque box in the following way: if it predicted that John will pick one box, it sets the amount to be \$1,000,000, and if it predicted that John will pick two boxes, it set the amount to \$0. The question is, should John pick two boxes, or should he pick one?

The debate on the Newcomb problem continues, but one aspect of it has been neglected. It has long been recognized that one's views on the Newcomb problem are somehow connected with, and motivated by, certain conditionals. So, Lewis expressed the basic one-boxing intuition as follows: '[One-boxers] are convinced by indicative conditionals: if I take one box I will be a millionaire, but if I take both boxes I will not.' (Lewis 1981a, 377). The proponents of two-boxing, by contrast, often appeal to subjunctives, as Lewis himself did: 'We [two-boxers] are convinced by counterfactual conditionals: If I took only one box, I would be poorer by a

thousand than I will be after taking both' (Lewis 1981a, 377). But, surprisingly, the nature of this connection has not been addressed head-on. How, precisely, do conditionals convince us of one or the other answer to the Newcomb problem? That's the question I want to pursue.

There are really two questions here: one is, which conditionals are we to appeal to, indicatives or subjunctives? The second, main, question is: *how* are we guided by conditionals toward one or the other answer to the Newcomb problem? Let me start with the first question.

Lewis, because he saw the contest between the one-boxer and the two-boxer as a contest between appeals to two different kinds of conditionals, thought that the 'debate is hopelessly deadlocked' (1981b, p. 5). But such pessimism is unjustified. Lewis appeared to think that whether one appeals to indicatives or to subjunctives is somehow a theoretical choice that one can make. But this choice is illusory. We in fact use indicative conditionals to think about what decisions to make: we ask, what will happen if I do X?, etc., -- nearly always in the indicative mood.³⁰ The suggestion that we might be wrong to use indicatives in deliberation carries the serious cost of positing a theory of practical deliberation according to which we systematically misuse the conceptual resources of English. What would recommend such a theory? I suspect that the only reason for Lewis' (and Gibbard's (1981b)) appeal to subjunctives in their discussion of the Newcomb problem is their *prior* commitment to an Adams theory of indicatives, and their belief that Adams indicatives would guide us to the wrong (one-boxing)

30. DeRose (2010) makes fundamentally the same point. But DeRose also holds that some indicatives are 'deliberationally useless' -- an issue I address below in §8.

There is also an important question here that I am not going to pursue: just what is the relation between future-oriented indicatives ('what will happen if I do A?'), and future less vivid subjunctives ('what would happen were I to do A?'). For one view of the matter, see DeRose 2010. For my purposes it is sufficient that we rightly use indicatives in a wide variety of decision situations, including the Newcomb problem.

answer to the Newcomb problem. I have argued above that indicatives do not have the Adams semantics; I will argue below that even if indicatives did have the Adams semantics, they would not guide us to one-boxing. But, even apart from these arguments, the cost of condemning so much of our ordinary practice is just too high. We should take our ordinary practice at face value and take indicatives to be the paradigmatic conditionals of rational deliberation. I will have nothing more to say about subjunctives, and turn to the role of indicatives in decision-theoretic reasoning.

So what *role* do indicative conditionals play in our thinking about what to do? *How* are we guided by conditionals toward an answer to the Newcomb problem? I will offer an answer to this question below, an answer that is non-partisan so far as the Newcomb problem is concerned. But while my main concern is to answer the *how* question, it will turn out that, with the help of the argument for Stalnaker's semantics from §2-6, my answer to the *how* question also yields an argument for two-boxing in the Newcomb problem (§9).

My proposal is simple. When we deliberate about whether A or B is the right course of action what we want to know is *whether we will be better off if we do A than if we do B*. In simple cases, being better off amounts to getting more of something -- more money, for example. In such cases, that is, in cases where *V-ing more is all that matters* (getting more money, or whatever), what we want to know is whether we will V more if we do A than if we do B. Of course, we are often not in a position to come to know this, and if we are not, we need to use other conceptual resources at our disposal to decide what to do. But, in the best-case scenario when we *can* come to know the relevant dominance indicative, I suggest that we find it of use because

we implicitly recognize the truth of the following decision-theoretic norm:

Dominance Norm (DN)

If the indicative dominance conditional 'if I do A, then I will V more than if I do B' is true in every world in the epistemic context of utterance, and if V-ing more is all that matters in the context, then option A is rationally preferable to option B.³¹

A couple of comments about DN, before we turn to its consequences.

i) DN is of course a kind of dominance principle -- a decision-theoretic principle that says that in some cases one course of action is preferable to another no matter what the relevant probabilities are. In particular, since I will eventually argue that DN can be used as part of an argument for two-boxing in the Newcomb problem, it is important to compare DN with the dominance principle sanctioned by causal decision theory:³²

Causal Dominance (CD)

Let $\{E_1, E_2, E_3, \dots\}$ be a partition of events that the decision maker regards as causally independent of her choice between A and B. If she weakly prefers A to B given E_j for every E_j , then she should weakly prefer A to B. Moreover, if any of these preferences is strict (and the associated event is nonnull), then she should strongly prefer A to B.

31. Why introduce the mention of *V-ing more* into DN, rather than appealing, straightforwardly, to being better off? Because the the notion of being better off is normative, and so may give rise to the suspicion that the truth of such a dominance indicative itself depends on which decision theory is right. Thus it may turn out that two parties to a decision-theoretic disagreement give different truth-values to the same better-off dominance conditional. This would prevent, or at least muddle the use of dominance conditionals to settle some decision-theoretic disagreements. By contrast, the formulation of DN given in the text distinguishes the normative assumption, that V-ing is all that matters, from the purely factual assumption, that we shall V more if we do A than if we do B.

32. For more on causal decision theory, see Joyce 1999.

(Joyce 1999, p. 151)

It is uncontroversial that CD leads to the two-boxing answer in the Newcomb problem. If we take {There is \$1,000,000 in the opaque box, There is \$0 in the opaque box} to be the partition, it is plausibly causally independent of one's choice (since the amount, whatever it is, had already been deposited in the opaque box). It then follows from CD that two-boxing is the right choice. But CD can hardly be used as part of an argument in favor of two boxing, because the two-boxing intuition in the Newcomb problem is itself supposed to be one of the main pieces of evidence in favor of CD.³³

But DN is importantly different from CD.³⁴ First of all, DN makes no mention of a partition of the space of states of nature or of causal independence, as CD does. Second, DN is not a recognized norm of causal or evidential decision theory, simply in virtue of the fact that standard formulations of decision theory make no mention of English indicative conditionals. Third, and this is crucial, the attractiveness of DN does not rest on any prior decision-theoretic commitments, or on judgements of rationality in contested scenarios (a charge that can be brought against CD). DN is plausible simply because it is plausible that there can be no decision situation in which the relevant dominance indicative is known to be true, and yet the course of action it recommends is the wrong course of action. DN can be plausibly considered to be constitutive of rational decision-making. So DN should be thought neutral on the

33. Pollock (forthcoming), p. 6, makes this same point explicitly. See *ibid.* for discussion of the relation between the Newcomb and the Smoking Lesion cases.

34. DN is of course also different from the unrestricted dominance principle that was employed in the original formulation of the Newcomb problem (i.e. the principle we get from CD by omitting the requirement of causal independence).

Newcomb problem.³⁵

ii) The natural default view is that a proposition P is known iff P is true in every world in the epistemic context of utterance. So it would be natural to formulate DN by appeal to knowledge of the relevant dominance conditional, rather than to its truth throughout C . But this is contested ground: the relation between knowledge and epistemic possibility is not clear. For example, Greco (forthcoming) argues that knowledge is compatible with epistemic probability less than 1, and so compatible with the falsehood of the proposition known in some worlds in C . Further, it may well be that truth throughout C is insufficient for knowledge. For these reasons I have formulated DN without appeal to knowledge. I will, however, purely for the sake of readability, continue to speak as though DN appeals to knowledge of the relevant dominance conditionals, with the understanding that, if the natural default view is false, a more precise formulation should always appeal to truth throughout C instead, and make no reference to knowledge.

iii) The comparison of DN with CD brings up another important point. CD, as I have given it in Joyce's formulation, is silent on what is the set of worlds over which the relevant partition is to be given. Clearly, that set of worlds is the set of possibilities relevant to decision-making.

35. One further question remains: what is the relation of DN and CD on the supposition that Stalnaker's semantics is the right semantics for indicatives? Clearly, since DN does not require causal independence, it has a wider application than CD, and does not follow from it. But does CD follow from DN, given Stalnaker's semantics? Perhaps it does. The question is: supposing that E_i is causally independent of A and B , and supposing that the actual world is an E_i world, does it follow that if A , then E_i , and if B , then E_i ? Does the indicative conditional and causal independence always go hand in hand? We certainly know that they usually do, but whether it is always so is unclear.

On the other hand, the formulation of DN appeals to C, the epistemic context of utterance. The notion of epistemic context comes from the work on the semantics of the epistemic modals, most notably the epistemic 'possibly', 'probably', and the indicative 'if'. So it is not obvious that the possibilities in C are precisely the possibilities that are decision-theoretically relevant. Without the assumption that C contains precisely the decision-theoretically relevant possibilities DN is clearly wrong. It is, however, quite plausible that C contains all and only possibilities relevant for decision-making: if it might be the case that P, then surely P is one of the possibilities that is relevant in decision-making, and if it is (epistemically) impossible that P, then surely the possibility that P is not decision-theoretically relevant. So the assumption that C contains precisely the decision-theoretically relevant possibilities has to be understood as part of the proposal I am putting forward.

iv) It is crucial that DN appeals to knowledge (i.e. truth throughout C), rather than *mere* probability of the relevant dominance conditional (i.e. probability < 1). DN manifestly fails if the dominance conditional is merely probable, because in that case the relative sizes of payoffs matter. To illustrate, imagine choosing between two lotteries. Lottery A pays \$1,000,000 with probability .1, and \$0 with probability .9. Lottery B pays \$1 with probability .9 and \$0 with probability .1. Now consider

(30) If I play lottery B I will win more than if I play lottery A.

Although one needs some further assumptions about the semantics of 'probably' to demonstrate the fact, it is plausible that (30) is probable (and assertible),

but not known to be true, given the description of the scenario. But clearly one ought to play lottery A, not lottery B. So the knowledge requirement cannot be relaxed.

I offer DN as the answer to the question raised earlier: how do conditionals guide us in reasoning about the Newcomb problem? What good is DN? DN is of use because it establishes a substantive connection between semantics and decision theory. In cases in which V-ing more is all that matters, DN connects the (known) truth of a *factual* claim (the relevant dominance indicative) with rational preference. In the case of the Newcomb problem it is clear that the normative assumption, that money is all that matters, is not in dispute. So, any semantics of indicatives that predicts that the one-boxing, or the two-boxing dominance indicative is known to be true commits us, *via* DN, to one-boxing, or to two-boxing.

To the extent that the debate over the Newcomb problem is a debate about what course of action conditionals recommend (as Lewis suggested), DN can help us formulate that debate precisely: as a disagreement over whether the right semantics of indicative conditionals commits us, *via* DN, to one- or two-boxing. In §9 I will look at the connection between two-boxing and Stalnaker's semantics, and one-boxing and Adams theories, in the light of DN. But before I turn to that, there is an objection to DN that needs to be considered.

8 Deliberationally useless indicatives

DeRose (2010) also holds, as I do, that indicative conditionals are rightly used in practical de-

liberation. But he thinks that there are cases when indicatives cannot be used to guide action - they are *deliberationally useless* (DeRose 2010, p. 20ff.). A conditional is deliberationally useless if it is known, and it recommends some course of action, but it would be irrational to heed the conditional's advice. The threat of deliberationally useless indicatives to my account is immediate: if some indicatives, in particular, if some dominance indicatives, are deliberationally useless, then DN falls, and with it the claimed connection between semantics and decision theory, at least in the strong form in which I suggest it. So we need to examine DeRose's case carefully. DeRose's scenario is as follows:³⁶

Risk It!

'Sly Pete is playing a new card game called Risk It! against Gullible Gus. Largely because your henchmen have been hovering about the game and helping him to cheat, the unscrupulous Pete has already won £1,000 from Gus as they move into the final round of the game. The final round of this game is quite simple. A special deck of 101 cards, numbered 0-100, is brought out, shuffled, and one card is dealt to each of the two players. After each player gets a chance to view his own card, but not his opponent's, the player who is leading going into the final round - in this case, Pete - gets to decide whether he wants to 'play' or 'quit'. If he decides to 'quit', then he simply keeps the money he has won before this final round - in this case, £1,000. If he instead decides to 'play', then his winnings are either doubled or cut to nothing depending on which player holds the higher card: both players show their card, and if the leader's (Pete's) is the higher card, the leader's winnings are doubled - in this case, to

36. DeRose's example is inspired by Gibbard's Sly Pete case (Gibbard 1981a).

£2,000. But if the leader decides to play, and his card is the lower one, he walks away with nothing.

In our first version of the story, your henchman Sigmund (the signaller) has seen what card Gus is holding, has signalled to Pete that Gus's card is 83, and has received Pete's return sign confirming that Pete got the message, and knows that Gus is holding 83. Sigmund does not know what card Pete is holding, and so does not know which player holds the higher card, but because he knows that Pete knows what both cards are, and because he is certain that Pete is not stupid enough to 'play' if his card is the lower one, it is clear that Sigmund knows that, and is in a position to report to you that:

(O) If Pete plays, he will win

Such information is helpful to you, because, we may suppose, you are making derivative bets on the results of Pete's game.' (DeRose 2010, p. 20-21)

But, DeRose points out,

But though Sigmund seems to know that, and seems in a position to report to you that, Pete will win if he plays, Pete cannot use this conditional that Sigmund knows in Pete's deliberation about whether or not to play. If Pete overhears Sigmund reporting to you that 'If Pete plays, he will win', it would be disastrous for Pete to reason as follows: 'Well, Sigmund seems to know that I'll win if I play; so, I should play.' And if Pete knows what Sigmund's grounds are for his claim, Pete will know not to reason in that disastrous way, if he is a competent consumer of indicative conditionals. This is a case where using a straightforward FDC [future-directed conditional, i.e. an indicative like (O)] as a conditional of deliberation leads to trouble: where the conditional would constitute bad advice if used in deliberation. (DeRose 2010, p. 21)

So, DeRose concludes that (O) is deliberately useless, and further notes that the cases where indicatives are deliberately useless include the Newcomb problem (I accept that 'if I choose one box, I will win a million', but this is 'bad advice', since I still should choose two boxes -- that's just the Lewis view). So, if matters are as DeRose describes, DN fails, and fails in precisely the case I am concerned to apply it to. However, I think DeRose's scenario does not show what it purports to show.

Let's grant that in the scenario Sigmund knows that if Pete plays, he will win. But is Sigmund's knowledge consistent with the possibility that Pete nevertheless goes on to play even though his card is the lower one -- by mistake, because perhaps his eyesight is poor and he misreads his own card, or intentionally, because he is suddenly overcome with remorse, and wants to make it up to poor Gus? If the answer is yes, then knowledge and truth throughout C can diverge. Call this the *weak* reading of Risk It! On the other hand, if the answer is no, then knowledge entails truth throughout C: Sigmund's knowledge rules out the possibility of Pete playing and losing. Call this the *strong* reading of Risk It!

Now, my response to the weak reading of Risk It! is just that if some worlds in C are worlds in which Pete plays and loses, then the relevant dominance indicative

(31) If Pete plays, he will win more money than if he does not.

will not be true throughout C -- it will be false at least in the worlds in which Pete plays and loses. But then DN does not apply, and Risk It! is not a counterexample to DN.

Now let's turn to the strong interpretation of DeRose's example. On the strong interpretation we are imagining that Sigmund's knowledge rules out the possibility of Pete playing and

losing. But now there is a different problem. What reason do we have for thinking that in the strong case the indicative (O), and the dominance indicative (31) recommend the *wrong* course of action? Why does DeRose, for instance, think that (O) recommends the wrong course of action in this case? Unfortunately, all DeRose says is that it would be 'disastrous' for Pete to use (O) to guide his actions. One would like to know what 'disastrous' means here. 'Disastrous', I take it, means that if Pete follows the advice of (O) (that is: play no matter what), he will do worse than if he follows a different strategy: to play if he has a higher card. But what would make us think that the first strategy is worse than the second? Presumably, that it is (epistemically) possible that the first strategy produces a worse outcome than the second -- that is, that there are cases where Pete has a lower card, and plays nevertheless (and so loses).³⁷ But if Sigmund knows, as per the strong interpretation, that there are no such cases, then the two strategies are in fact equally good in the strong case: if it is not possible to play and lose, Pete might as well play.³⁸

So my response to the DeRose worry in the strong case is that Pete can indeed trust the advice given by (O). And so again there is no counter-example to DN. Now let's come back to the connection between decision theory and semantics established by DN.

37. One might propose that the use of (O) is disastrous if Pete *would* lose if he *were* to play. But it is not clear why Pete should be concerned with the advice being disastrous in this way. If it is epistemically impossible for him to play and lose -- why should it be of concern how he *would* do were he to play?

38. Admittedly, it would be rather strange for Pete to overhear Sigmund's utterance of (O), and resolve to play. But the strangeness of this course of events is all due to the unrealistic nature of the strong reading: if Pete knows (O), he knows that he will not play and lose, even before he looks at his own cards.

9 The semantics of indicatives and the Newcomb problem

What can be learned about the Newcomb problem given the results in semantics from §2-6, and the foregoing defense of DN? Let's take it for granted that the two-boxers and the one-boxers are both guided by indicative conditionals. Let's further assume that they are both guided by indicative conditionals with the help of DN. Here there are two interesting connections, one between Stalnaker's semantics and two-boxing, the other between one-boxing and Adams theories.

Let's start with Stalnaker. The Newcomb problem is a variably dominant scenario: the amount of money in both boxes dominates the amount of money in one (since it is epistemically impossible for the two boxes to contain an amount less than or equal to the amount of money in the opaque box). The two relevant dominance conditionals are

(11) If I choose one box, I will win more than if I choose two.

(12) If I choose two boxes, I will win more than if I choose one.

The reasoning here is perfectly parallel to the reasoning given for T1 and (10) in §5. Since the choice of one or two boxes does not causally influence the amounts of money in either box, (12), the two-boxing dominance conditional, is true, and known to be true, in the Newcomb case. So, by DN, Stalnaker's semantics commits us to two-boxing.

This result is of course not surprising: although, as I have suggested, the precise way in which conditionals guide decision-making has not so far been made clear, it has always been

agreed on all sides that subjunctive conditionals recommend two-boxing, and Stalnaker's semantics for indicative conditionals makes them very similar to the Stalnaker-Lewis subjunctives. What is new in my argument is the status of DN. Gibbard & Harper's (1981b) causal decision theory is useless as a means of settling the dispute between the one-boxer and the two-boxer, because it rests, from the one-boxer's point of view, on question-begging assumptions about the import of causal relationships.³⁹ DN, by contrast, is neutral, and so is an appropriate ground to conduct the dispute.

What about Adams theories? Does the present discussion bear out Lewis' claim that one-boxers are guided by indicative conditionals, provided that we grant Lewis' unspoken assumption these in turn get the Adams semantics?⁴⁰ Here matters are more complicated. Let's grant the Adams theorist that the one boxing dominant conditional

(11) If I pick one box I will win more than if I pick two boxes.

is assertible.⁴¹ But since we cannot assign a truth-value to (11), DN, as I have formulated above, does not apply. Furthermore, the natural variation on DN:

Dominance Norm -- Adams (DNA)

If the indicative dominance conditional 'if I do A, then I will V more than if I do B' is *assertible*, and if V-ing more is all that matters in the context, then option A is rational-

39. For a similar observation, compare Pollock (forthcoming), p. 5.

40. I am ignoring here the differences between Lewis' (1976) sophisticated material implication analysis of indicatives and the more common Edgington-style view.

41. One can argue as follows: If I pick one box, I will win \$1,000,000; if I pick two boxes, I will win \$1000, therefore, if I pick one box, I will win more than if I pick two boxes. The premises of this argument are assertible on Adams semantics because the corresponding conditional probabilities are high. Further, the argument is valid. So, the conclusion, the one-boxing dominance indicative, is also assertible.

ly preferable to option B.

clearly does not work, for the reason given in §7. So, if lottery A pays \$1,000,000 with probability .1, and \$0 with probability .9, and lottery B pays \$1 with probability .9 and \$0 with probability .1, then

(30) If I play lottery B I will win more than if I play lottery A.

is assertible, and so DNA recommends playing lottery B. But clearly one should play lottery A. So DNA is wrong.⁴²

There do not seem to be any plausible DN-like principles that would connect with (11) to yield the recommendation of one-boxing. So either it is a mistake to think that Adams theories of indicatives favor one-boxing, or they do favor one-boxing, but with the help of some other principle besides DN. I am inclined to think that the first possibility is true, but the question needs further study.

To sum up: if DN is the way to understand how indicatives are involved in practical deliberation, then we have two results, one positive and one negative. The positive argument appeals to DN and the argument for Stalnaker's semantics from §2-6: if DN is right, then if Stalnaker-type semantics is right, we should two-box. The negative result is that Adams indicatives do not recommend one-boxing via DN, or any similar principle.

42. On the other hand, if we try to fix DN* by stipulating that the dominance conditional is to have assertibility 1, the resulting principle will no longer apply in the Newcomb case, since (11) plausibly does not have assertibility 1.

10 Conclusion

The argument of this paper can now be put succinctly. The dominance indicative (10) is true in the key variably dominant scenario T1. But only one type of theory of indicatives can predict this: a Stalnaker-type semantics. Further, we should accept DN, the principle that tells us that, in the right circumstances, the known truth of the relevant dominance indicative recommending option A entails that option A is in fact the rational course of action. But DN in combination with a Stalnaker-type semantics entails that two-boxing is the rational choice in the Newcomb problem.

APPENDIX

The two main theories of the comparative are the degree-based and the interval-based approaches. Here I will lay out the predictions of each theory for (10) in combination with Kratzer's semantics and, for comparison, with Stalnaker's theory. The end result is the same as that reached above: Kratzer's theory makes the wrong predictions either for T1 or T2, and Stalnaker's semantics gives the right predictions, no matter which theory of the comparative is used to derive them.

Let's start with the standard degree-based account of comparatives.⁴³ On the standard account,

(32) Mary is taller than Paul.

compares two *sets* of heights -- those which Mary meets or exceeds, and those that Paul meets or exceeds -- and is true just in case some specified relation holds between these two sets of degrees. It is natural to think that the relation is that the maximal element in the first set is greater than the maximal element in the second set.

The standard semantics consists of a semantics for gradable expressions, and a semantics for the comparative morpheme '-er'. It can be summarized as follows (the following exposition follows Heim (2006)).

(33) $\llbracket \text{tall} \rrbracket = \lambda d_d. \lambda x_e. x\text{'s height} \geq d$

(34) $\llbracket \text{-er} \rrbracket = \lambda P_{\langle d, t \rangle}. \lambda Q_{\langle d, t \rangle}. \max(Q) > \max(P)$

43. For more on standard accounts see Stechow 1984, Heim, 2006, Kennedy 2007, Schwarzschild 2008.

So the Logical Form (LF) for (32) will look like this:

$$(35) \quad [-\text{er than wh}_1 [\text{Paul is } t_1 \text{ tall}]]_2 [\text{Mary is } t_2 \text{ tall}]^{44}$$

which leads to the following truth-conditions:

$$(36) \quad \max(\{d: \text{Mary is at least } d\text{-tall}\}) > \max(\{d: \text{Paul is at least } d\text{-tall}\})$$

which is true just in case Mary's height is greater than Paul's, as desired.

Now let us see what the standard account predicts for dominance indicatives on Kratzer's theory. Let us take

$$(10) \quad \text{If I drink the contents of cup A, I will ingest more than if I drink the contents of cup B.}$$

The LF will be as follows

$$(37) \quad [-\text{er than wh}_1 [\text{if I drink the contents of cup B, I will ingest } t_1]]_2 [\text{If I drink the contents of cup A, I will ingest } t_2]$$

which gives the following truth-conditions:

$$(38) \quad \max(\{d: \text{if I drink the contents of cup A, I will ingest at least } d\text{-much}\}) > \max(\{d: \text{if I drink the contents of cup B, I will ingest at least } d\text{-much}\})$$

which, if we now add Kratzer's semantics for the indicative conditional, becomes

$$(39) \quad \max(\{d: \forall w (\text{I drink from cup A in } w \supset \text{I will ingest at least } d\text{-much in } w)\}) > \max(\{d: \forall w (\text{I drink from cup B in } w \supset \text{I will ingest at least } d\text{-much in } w)\})$$

So, according to the standard account of comparatives in combination with Kratzer's se-

44. Here the -er phrase is of type $\langle d, t \rangle$, and so moves because of a type-mismatch, since 'tall' needs an argument of type $\langle d \rangle$. It is a standard assumption that the 'than'-clause involves a vacuous wh-operator that induces lambda-abstraction (see Heim 2006 for the details).

mantics, (10) is true just in case the maximum amount d such that 'I drink the contents of cup A \supset I will ingest at least d -much' is true in all $w \in C$ is greater than the maximum amount d' such that 'I drink the contents of cup B \supset I will ingest at least d' -much' is true in all $w \in C$. In other words, according to (39), (10) is true just in case the *minimum* I drink in all the worlds in which I drink from cup A is greater than the *minimum* I drink in all the worlds I drink from cup B. But (39), as can be easily seen, gives the wrong predictions. So consider the following:

Tea-party 2

Two cups of tea are on the table -- cup A, and cup B. We know that cup A contains at least 10ml of tea, and cup B at least 1ml of tea, and nothing else of relevance (in particular we do not know that cup A contains more tea than cup B).

Let's evaluate (10) in T2, assuming that (39) gives the right truth-conditions. Clearly, $\max(\{d: \text{if I drink from cup A, I will ingest at least } d\text{-much}\}) = 10\text{ml}$ (because for all $d > 10\text{ml}$ it is false that in all the worlds in which I drink from cup A, I ingest at least d -much) and $\max(\{d: \text{if I drink from cup B, I will ingest at least } d\text{-much}\}) = 1\text{ml}$. So the standard semantics for comparatives, coupled with Kratzer's semantics for indicatives predict that (10) is true in T2. But this is the wrong result.

I think it is plausible that (37) is the LF of (10) because it is plausible that (10) is equivalent to

(40) The amount I will ingest if I drink from cup A is greater than the amount I will ingest if I drink from cup B

-- and in (40) the comparative clearly out-scopes both conditionals. But note that when the -

er-phrase moves, it has two landing sites available, and the other possibility is

- (41) [If I drink the contents of cup A [-er than wh₁ [if I drink the contents of cup B, I will ingest t₁]]₂ [I will ingest t₂]]

Combining (41) with Kratzer's semantics raises a further issue. When I introduced Kratzer's view in §71, I introduced the simple version of the theory for *straight* indicatives -- indicative conditionals without modals in the consequent. But (41) is not a straight indicative, because its consequent contains another conditional. The question is whether the consequent of (41) is to be interpreted against the old modal base C, or against an updated modal base $C \cap \llbracket \text{I drink the contents of cup A} \rrbracket^C$. I think the second option is preferable in general (that is the route suggested by Gillies 2009), but let me spell out the consequences of taking both options.

On the first option, (41) combined with Kratzer's semantics yields:

- (42) $\forall w (\text{I drink from cup A in } w \supset (\max(\{d: \text{I will ingest at least } d\text{-much in } w\}) > \max(\{d: \forall w (\text{I drink from cup B in } w \supset \text{I will ingest at least } d\text{-much in } w\})))$

(42) makes (10) true just in case the minimum amount I drink in the worlds in which I drink from cup A is greater than the minimum amount I drink in the worlds in which I drink from cup B. These truth-conditions, as we have seen give the wrong verdict in the case of T2.

If the consequent of an indicative conditional is evaluated in a shifted context, $C \cap \llbracket \text{I drink the contents of cup A} \rrbracket^C$, then (41) gives

- (43) $\forall w (\text{I drink from cup A in } w \supset (\max(\{d: \text{I will ingest at least } d\text{-much in } w\}) > \max(\{d: \forall w (\text{I drink from cup A in } w \text{ and I drink from cup B in } w \supset \text{I will ingest at least } d\text{-much in } w\})))$

But note that (43) is trivially true, given that C does not contain any worlds in which I drink from both cups. In sum, (41) does not fare better than (37).

Both Heim (2006) and Schwarzschild&Wilkinson (2002) have suggested that the standard account of comparatives gives the wrong predictions in cases like the following:

(44) John is taller than every girl in his class.

The standard semantics predicts that (44) is true iff

(45) $\max(\{d: \text{every girl in his class is at least } d\text{-tall}\}) > \max(\{d: \text{John is at least } d\text{-tall}\})$

that is, iff John is taller than the *shortest* girl in his class, which is clearly the wrong result.

The standard semantics predicts a *greater-than-minimum* reading where a *greater-than-maximum* reading is called for. Since Kratzer's semantics for indicatives makes them essentially universal quantifiers (over worlds in the epistemic context), the difficulty with T2 illustrated above can be seen as due to the same kind of failure as the failure of (45), and so be blamed on the degree semantics of comparatives. So we need to see how Kratzer's view fares on other approaches to comparative semantics. To remedy the problem with (45) Heim proposed an interval-based semantics of comparatives. Let me sketch it briefly.⁴⁵

First, the lexical entry for the adjective:

(46) $\llbracket \text{tall} \rrbracket = \lambda D_{\langle d, t \rangle}. \lambda x_e. x\text{'s height} \in D$

⁴⁵ I am here presenting the simpler of the two theories that Heim discusses in 2006. The simpler account makes the wrong predictions in some of the cases that the standard degree-based semantics gets right. (E.g. 'I may eat more than I have to'; thanks to Will Starr for the example). But this will not matter for our purposes, since Heim's more complex Π -operator theory gives the same results in our cases.

Here, D is a set of degrees. Next, the lexical entry for the comparative morpheme '-er':

$$(47) \quad \llbracket \text{-er} \rrbracket = \lambda d_d. \lambda d'_d. d' > d$$

We can see how these lexical entries work if we look at the example Heim discusses in (2006):

$$(48) \quad \text{The desk is wider than every couch is long.}$$

The LF is:

$$(49) \quad [\text{wh}_1 [\text{every couch is } t_1 \text{ long}]]_2 [\text{the desk is } [-\text{er than } t_2] \text{ wide}]^{46}$$

which leads to the following derivation:

$$\begin{aligned} (50) \quad & [\lambda D_{\langle d, t \rangle}. \forall x_e [\text{couch}(x) \rightarrow x \text{'s length} \in D]] (\lambda d_d. \text{the desk's width} \in [\lambda d'_d. d' > d]) \\ &= [\lambda D. \forall x [\text{couch}(x) \rightarrow x \text{'s length} \in D]] (\lambda d. \text{the desk's width} > d) \\ &= \forall x [\text{couch}(x) \rightarrow x \text{'s length} \in [\lambda d. \text{the desk's width} > d]] \\ &= \forall x [\text{couch}(x) \rightarrow \text{the desk's width} > x \text{'s length}] \end{aligned}$$

So Heim's account predicts, correctly, that (48) is true just in case the desk's width is greater than *every* couch's length (and similarly for (44)). Now, let's replicate this for a Kratzer dominance conditional. Again, let's come back to our example:

$$(10) \quad \text{If I drink the contents of cup A, I will ingest more than if I drink the contents of cup B.}$$

The LF will be as follows:

$$(51) \quad [\lambda D_{\langle d, t \rangle}. \text{if I drink from cup B, I will ingest an amount } d \in D] (\lambda d_d. \text{if I drink from cup A, I$$

46. Here, the -er phrase has the right type, $\langle d, t \rangle$, but the than-phrase, which is of type $\langle dt, t \rangle$ has to move for type reasons.

$$\begin{aligned}
& \text{will ingest an amount } \in [\lambda d'_{d,d'} > d]) \\
& = [\lambda D_{<d,l>}. \text{ if I drink from cup B, I will ingest an amount } d \in D] (\lambda d_d. \text{ if I drink from cup A,} \\
& \text{I will ingest an amount } > d) \\
& = [\lambda D_{<d,l>}. \forall w (\text{I drink from cup B in } w \supset \text{I will ingest an amount } d \in D \text{ in } w)] (\lambda d_d. \forall w (\text{I} \\
& \text{drink from cup A in } w \supset \text{I will ingest an amount } > d \text{ in } w))
\end{aligned}$$

In other words, Kratzer's semantics in combination with the Heim interval semantics for comparatives predicts that (10) is true just in case the *minimum* amount I drink in all the worlds in which I drink from cup A is greater than the *maximum* amount I drink in all the worlds in which I drink from cup B (if there is a maximum -- if there is not, (10) will be false). So (51) predicts that (10) will be false in T2.

But (51) gives the wrong result for T1. In T1, the denotation of $[\lambda D_{<d,l>}. \forall w (\text{I drink from cup B in } w \supset \text{I will ingest an amount } d \in D \text{ in } w)]$ is $\{D: 1\text{ml} \in D \text{ and } 3\text{ml} \in D\}$. The denotation of $(\lambda d_d. \forall w (\text{I drink from cup A in } w \supset \text{I will ingest an amount } > d \text{ in } w))$ will be $[0, 2)$ (the minimum I drink in the A-worlds is 2ml). But $[0, 2) \notin \{D: 1\text{ml} \in D \text{ and } 3\text{ml} \in D\}$, so according to (51), (10) is false in T1. That is the wrong result.

To sum up: when combined with either of the two main approaches to comparatives, Kratzer's semantics predicts that (10) is false in T1 or true in T2. This is the wrong result.

It is not surprising that Stalnaker's semantics gives the right predictions for T1 and T2 on either theory of the comparative.

Recall that the degree approach gives the following LF:

- (37) [-er than wh₁ [if I drink the contents of cup B, I will ingest t₁]]₂ [If I drink the contents of cup A, I will ingest t₂]

which gives rise to the following truth-conditions:

- (38) $\max(\{d: \text{if I drink the contents of cup A, I will ingest at least } d\text{-much}\}) > \max(\{d: \text{if I drink the contents of cup B, I will ingest at least } d\text{-much}\})$

And Heim's interval semantics gives the following LF:

- (52) $[\lambda D_{<d, \triangleright}. \text{if I drink from cup B, I will ingest an amount } d \in D](\lambda d_d. \text{if I drink from cup A, I will ingest an amount } \in [\lambda d'_d. d' > d])$
 $= [\lambda D_{<d, \triangleright}. \text{if I drink from cup B, I will ingest an amount } d \in D](\lambda d_d. \text{if I drink from cup A, I will ingest an amount } > d)$

Both (38) and (52), in combination with Stalnaker's semantics, are true at a world w in C just in case the amount I drink in the closest world to w in which I drink from cup A is greater than the amount I drink in the closest world to w in which I drink from cup B. This gives the right verdict for T1: (10) comes out true at every world in C . And the prediction for T2 is right, too: (10) will come out true in some worlds in C , namely in those in which the amount in cup A is greater than the amount in cup B, and false in others, namely in those in which the amount in cup A is less than or equal to the amount in cup B. So, while true in some worlds in C , (10) will not be assertible in T2.

Chapter 3

Whether-Conditionals

1 Introduction

In this paper I look at *indicative nested whether-conditionals*, sentences like the following:

(53) If I run the marathon, I'll lose whether I take performance-enhancing drugs or not.

I will argue that every theory of indicative conditionals that survives McGee's argument against *modus ponens* is inconsistent with the apparently deviant behavior of a particular species of nested whether-conditionals that I will call the *singular* nested whether-conditionals. The goal is not to demolish the old theories and then build a new semantics for indicatives from scratch, but to discover the mechanism by which we get the deviant readings, a mechanism that could be incorporated into some existing theory. I will suggest that the mechanism at issue is a peculiar kind of presupposition accommodation. Besides describing and explaining a previously unobserved phenomenon, there are two pay-offs. First, the mechanism I propose is unusual, and so is of independent theoretical interest. Second, it turns out that the mechanism I propose only supplies the desired explanation when combined with one particular theory of indicative conditionals -- Heim's (1992) account. So singular whether-condition-

als give us an argument in favor of Heim's semantics (augmented with the special mechanism that I propose).

2 McGee's argument against modus ponens

In §4, I will present a scenario involving nested whether-conditionals that every reasonable theory of indicative conditionals gets wrong (call a theory *reasonable* if it survives McGee's argument against *modus ponens*). The nature of the problem is best seen by contrasting singular nested whether-conditionals with *or-to-if* conditionals, conditionals of the form \neg if $\varphi \vee \theta$, then if $\sim\varphi$, then θ . Briefly put, or-to-if conditionals make it plausible that 'if' shifts the epistemic context (that's the *normal* behavior), while the deviant nested whether-conditionals require that 'if' does not shift the epistemic context. The *central problem* is to explain how the two behaviors can co-exist. Our first goal is to formulate the central problem.

Let's start with or-to-ifs. To appreciate the role of or-to-ifs it is best to review McGee's argument against *modus ponens*. An added benefit is that McGee's argument will rule out Stalnaker's semantics for indicatives, which happens to be the only theory on the market that predicts the deviant readings.

Let's start with Stalnaker's semantics. Here is an intuitive formulation, to be improved on in §3:

Stalnaker's semantics (rough formulation)⁴⁷

‘If φ , then ψ ’ is true at w iff the closest φ -world to w is a ψ -world.

(here the closeness relation, also known as the *similarity relation*, is the relation familiar from Stalnaker's and Lewis' analysis of counterfactuals)

Now, consider this familiar pattern of entailment:

Unrestricted Modus Ponens (UMP)

$\vdash \varphi, \vdash \text{if } \varphi, \text{ then } \psi \vdash \vdash \psi$. (for all φ, ψ)

Stalnaker's semantics validates UMP: if φ is true at the actual world, α , then the closest φ -world to α is α , and so ‘if φ , then ψ ’ is true just in case α is a ψ -world.

In his (1985), McGee presented a counter-example to UMP, and therefore to Stalnaker's semantics. McGee's central example is as follows. We have two Republican candidates, Reagan and Anderson, and one Democrat -- Carter. Reagan is overwhelmingly likely to win, with Carter running in a rather distant second place, and Anderson a very distant third -- he barely has a chance. Now, we argue (before the election):

The 1980 Election

(54) If a Republican wins, then if it's not Reagan who wins, it will be Anderson.

(55) A Republican will win.

(56) So, if it's not Reagan who wins, it will be Anderson.

We can grant that the conclusion is false: if Reagan does not win, Carter will. Further, (54) is

47. Let me stress that I am talking only about indicatives in this paper, so ‘Stalnaker's semantics’ always refers to his semantics of indicatives.

surely true. What about (55)? Well, by hypothesis, we do not know for sure that it is true, but we can consistently suppose that it is.⁴⁸ So, we have a counterexample to UMP, and to Stalnaker's semantics. Note that Stalnaker's semantics predicts that (54) will be false in the Election. So one might object, on behalf of Stalnaker's view, that the Election is in fact not a counter-example to UMP. But this only points to the fact that what is really doing the work in the Election is the claim that (54) is true, a claim Stalnaker has to deny. So McGee's argument stands.

It is crucial that the Election is not merely a counter-example to *modus ponens*; it also presents an explanatory challenge. McGee observes that the pattern instantiated by (54) is general, that is to say, that nested conditionals like (54) seem to be logically true. Call conditionals of the form 'if φ then θ , then if $\sim\varphi$, then θ ' *or-to-if conditionals*.⁴⁹ Now, this fact demands explanation. Of course, some semantic theories of conditionals may just deliver this result, and this would count in their favor. But what one would like is some sort of explanation: some general feature that is responsible for the fact that or-to-if conditionals are logically true. McGee provides one answer to this question. He remarks: 'It appears, from looking at examples, that the law of exportation, 'if φ and ψ , then θ ' entails 'if φ , then if ψ , then θ ', is a feature of English usage', and adds in a footnote that importation ('if φ , then if ψ , then θ ' entails 'if φ and ψ , then θ ') seems valid as well (1985, p. 465). So McGee, in effect, argues that the best explanation for the logical truth of or-to-if conditionals is the principle of Import-Export: 'if φ , then

48. This is a necessary feature of the example. If we *knew* that a Republican will win, the conclusion would plausibly be true, and not false, as intended.

49. The logical truth of the or-to-if conditional is to be distinguished from the validity of what is often called the or-to-if inference: φ or $\psi \models$ if not φ , then ψ . The or-to-if inference may well be invalid.

if ψ , then θ \models 'if φ and ψ , then θ '.

However, there is clearly a jump between our acceptance of 'if $\varphi \vee \theta$, then if $\sim\varphi$, then θ ' as a logical truth, and our acceptance of Import-Export. To be sure, Import-Export entails 'if $\varphi \vee \theta$, then if $\sim\varphi$, then θ ', but it is not obvious that the reverse entailment holds. The semantics of indicatives Heim proposed in (1992) is the only theory that accepts 'if $\varphi \vee \theta$, then if $\sim\varphi$, then θ ' as a logical truth, and yet fails to validate Import-Export. Heim's theory is of interest in two ways. First, it will help us formulate precisely what is to be learned from McGee's argument -- to diagnose the normal behavior of 'if'. Second, it is the theory I will ultimately defend. Let me sketch it briefly.

3 Heim's semantics

First, some semantic preliminaries. We shall be concerned with evaluating utterances of conditionals. I will use two parameters of evaluation: a world of evaluation, w , and an epistemic context, c , which I take to be the set of propositions known in the context of utterance.⁵⁰ Sometimes, it will also be helpful to speak of *the context of evaluation*, the pair $\{w, c\}$, consisting of a world of evaluation and an epistemic context.

Let $\text{sim}(w, \varphi)$ be the *closeness function*: for each world w and proposition φ , it returns a world w' such that w' is the closest φ -world to w . This definition incorporates the uniqueness assumption, the assumption that for each w , there is a unique closest φ -world w' . This simpli-

50. One can think of the epistemic context as a set of propositions, or as a set of worlds. It will be crucial to the proposal I develop in §7 that we think of the epistemic context in the first way.

fies exposition, and is harmless in the context of the present investigation. Note that by definition the closeness function is *unrestricted*: *sim* searches for the closest φ -world throughout the universe of worlds. If we want *sim* to search for a closest φ -world within some context *c*, we can do that by asking for the value of $\text{sim}(w, c+\varphi)$, where ' $c+\varphi$ ' is the union $\text{cu}\{\varphi\}$.⁵¹ So one can read ' $\text{sim}(w, c+\varphi)$ ' as 'the closest φ -world in *c* to *w*'. It is sometimes helpful to speak of *the set of worlds in c* -- we shall understand by that the set of worlds in which all the propositions in *c* are true; in this case $c+\varphi$ can be thought of as the intersection of the worlds in which all the propositions in *c* are true with the set of worlds in which φ is true.

Now let us turn to conditionals. In general, it is helpful to think of 'if' as a *context shifter*: that is, it is helpful to describe the evaluation of 'if φ , then ψ ' as proceeding in two stages: first, 'if φ ' changes the current context of evaluation in some way (by changing the epistemic context *c*, or the world of evaluation, or both); next, ' ψ ' is evaluated in the new, *derived* or *local*, context. Now we can formulate Heim's semantics (I give Stalnaker's view side-by-side just to illustrate the differences).

Stalnaker

An indicative 'if φ , then ψ ' is true in *w*, *c*, iff ' ψ ' is true in $\text{sim}(w, c+\varphi)$, *c*.

Heim

An indicative 'if φ , then ψ ' is true in *w*, *c*, iff ' ψ ' is true in $\text{sim}(w, c+\varphi)$, $c+\varphi$.⁵²

Both theories shift the world of evaluation from *w* to $\text{sim}(w, c+\varphi)$. The crucial difference be-

51. Since $\text{cu}\{\varphi\}$ is not a proposition, but a set of propositions, strictly speaking, $\text{sim}(w, c+\varphi)$ returns the closest world *w'* in which *all the propositions* in $c+\varphi$ are true.

52. This is a static version based on Heim's dynamic formulation: $c+\text{if } \varphi, \text{ then } \psi = \{w \in c : \text{sim}(w, c+\varphi)+\psi = \text{sim}(w, c+\varphi)\}$ (Heim 1992, p 196)

tween the two is that, according to Stalnaker, the consequent is evaluated in the original epistemic context, while according to Heim, the consequent is evaluated in a derived epistemic context, $c+\varphi$. As a result, the two theories agree on all cases where the consequent has no modally sensitive material. But if the consequent is itself a conditional, it is modally sensitive. So, crucially, the two theories diverge on or-to-if conditionals. Let me quickly illustrate the divergence.⁵³

Stalnaker starts evaluating $\ulcorner \text{if } \varphi \vee \theta, \text{ then if } \sim\varphi, \text{ then } \theta \urcorner$ by shifting the world, but not the context of evaluation: the or-to-if is true if $\ulcorner \text{if } \sim\varphi, \text{ then } \theta \urcorner$ is true in $\text{sim}(w, c+\varphi \vee \theta), c$. Then the nested antecedent shifts the world of evaluation once again: the or-to-if conditional is true iff $\ulcorner \theta \urcorner$ is true in $\text{sim}(\text{sim}(w, c+\varphi \vee \theta), c+\sim\varphi), c$. Naturally, on Stalnaker's semantics or-to-ifs can easily be false, as (54) demonstrates.

Now, turn to Heim's predictions for or-to-ifs. Start with $\ulcorner \text{if } \varphi \vee \theta, \text{ then if } \sim\varphi, \text{ then } \theta \urcorner$. We begin evaluating the or-to-if by shifting from the context w, c to the context $\text{sim}(w, c+\varphi \vee \theta), c+\varphi \vee \theta$. But now, since every world in the new context is either a φ -world or a θ -world, it is clear that $\ulcorner \text{if } \sim\varphi, \text{ then } \theta \urcorner$ will be true in every world in the new context. So or-to-if conditionals come out logically true. What is really doing the work in this reasoning is the assumption that the indicative 'if' shifts the epistemic context by updating on the antecedent.

I have suggested in the last section that McGee's argument against *modus ponens* gives rise

53. Here are the predictions for a generic right-nested conditional 'if φ , then if ψ, θ ':

Stalnaker: An indicative 'if φ , then if ψ, θ ' is true in w, c , iff 'if ψ, θ ' is true in $\text{sim}(w, c+\varphi), c$, iff ' θ ' is true in $\text{sim}(\text{sim}(w, c+\varphi), c+\psi), c$

Heim: An indicative 'if φ , then if ψ, θ ' is true in w, c , iff 'if ψ, θ ' is true in $\text{sim}(w, c+\varphi), c+\varphi$ iff ' θ ' is true in $\text{sim}(\text{sim}(w, c+\varphi), c+\varphi+\psi), c+\varphi+\psi$

to an explanatory challenge -- the challenge to explain the triviality of or-to-ifs. We have seen that McGee answers this challenge with Import-Export. Heim offers a different diagnosis: the epistemic context-shifting by the antecedent. But note that any theory committed to Import-Export effectively treats 'if' as if it shifted the epistemic context: if $\varphi \rightarrow (\psi \rightarrow \theta)$ is equivalent to $(\varphi \& \psi) \rightarrow \theta$, and the truth of $(\varphi \& \psi) \rightarrow \theta$ depends only on what happens in $(\varphi \& \psi)$ -worlds, then the truth of $\varphi \rightarrow (\psi \rightarrow \theta)$ can only depend on the truth of $(\psi \rightarrow \theta)$ in φ -worlds. Further, Heim's semantics invalidates Import-Export.⁵⁴ Heim's answer to the explanatory challenge is more conservative than McGee's. So it is plausible to think that the minimal lesson of or-to-if conditionals is that 'if' shifts epistemic context (by updating the original context with the antecedent). Epistemic context-shifting is the normal behavior of 'if's, and is clearly evident from many other examples besides or-to-ifs. For example, it is clearly the right explanation for probability conditionals: 'if φ , then probably ψ '.

,

4 Whether-conditionals

The lesson of or-to-ifs is one half of the central problem. In or-to-ifs 'if' behaves normally: it shifts the epistemic context. Now let's turn to the other half -- to the deviant behavior of 'if' in the singular whether conditionals. Consider the following situation:

Pass/Fail

⁵⁴. Here is a counter-model :

w1: $\sim\varphi$ w2: $\varphi, \psi, \sim\theta$ w3: $\varphi, \sim\psi$ w4: φ, ψ, θ

Suppose further that $\text{sim}(w1, \varphi \& \psi) = w2$. Then $(\varphi \& \psi) \rightarrow \theta$ is false in w1. But suppose further that $\text{sim}(w1, \varphi) = w3$, and $\text{sim}(w3, \varphi \& \psi) = w4$. Then $\varphi \rightarrow (\psi \rightarrow \theta)$ is true in w1.

Ivan's agnostic friend Joe is preparing for an exam, but fears that ordinary means might not be enough. He is trying to decide whether or not to pray, asking God to help him pass the exam. Joe asks Ivan for advice: should he pray, or not?

Ivan, an atheist, responds:

Ivan's argument

(57) If you pass, you'll pass whether you pray or not.

(58) If you fail, you'll fail whether you pray or not.

So, you shouldn't pray.⁵⁵

Let's start with three initial observations:

- i) Ivan's argument seems valid. (of course, that means: valid in the context, with the natural assumptions about utilities of things, etc.)
- ii) since the conclusion is not a logical truth, the premises better be informative.
- iii) Even quite apart from the argument, the premises clearly are informative.⁵⁶

55. The argument has a long history. Cf. Cicero, On Fate:

Nor will we be blocked by the so called 'Lazy Argument' (the *argos logos*, as the philosophers entitle it). If we gave in to it, we would do nothing whatever in life. They pose it as follows: 'If it is your fate to recover from this illness, you will recover, regardless of whether or not you call the doctor. Likewise, if it is your fate not to recover from this illness, you will not recover, regardless of whether or not you call the doctor. And one or the other is your fate. Therefore it is pointless to call the doctor...' (Long & Sedley, *The Hellenistic Philosophers*, 55S, p. 339).

Of course, there are important differences between Ivan's argument and the lazy argument, but investigation of this does not belong in this paper.

56. Pass/Fail appeals to our dynamic intuitions, intuitions about informativeness. But things are the same if we ask instead about our static intuitions -- intuitions about what is and is not true in a given context. So consider also Pass/Fail*: Pass/Fail with the stipulation that in fact there is a god who rewards Joe's prayers. Then the initial observation is: (57) and (58) are false in Pass/Fail*, (or at least unassertable).

So far, we have noted only that (57) and (58) seem informative in Pass/Fail, but what do (57) and (58) really mean? Suppose Joe accepts Ivan's assertion of (57) and (58) -- what does he learn? The most intuitive answer appeals to causal influence: (57) and (58) are informative, we would like to say, because they inform Joe that praying has no causal influence on his passing.

While I think that this answer is correct so far as it goes, one can do better. Suppose that we know that there is one and only one god, but that we are not sure about his character. One possibility is that our god is a god who *rewards prayer*.⁵⁷ Another possibility is that our god is a god who *punishes non-prayer*. These two are obviously distinct: we can conceive a god who interferes only when someone prays, and also a god who interferes only when someone does not pray. Somewhat less intuitively, we can also imagine a god who rewards non-prayer (a god who rewards self-reliance, so to speak). Likewise, we can imagine a particularly malicious god who punishes prayer. The four possibilities are of course not all mutually exclusive, although some combinations are peculiar -- such as the possibility that our god rewards both prayer and non-prayer. Now we can ask: which of these four possibilities does Ivan's assertion of (57) rule out?

I think the answer is pretty clear: it rules out all four possibilities. For example, suppose that it were a possibility that our god rewards non-prayer. Then, in the worlds in which Joe prays and fails he would have done better by not praying. Intuitively, the truth of (57) rules these worlds out. And another case: suppose that in the actual world, Joe prays and passes.

57. For the record, note that there are at least two ways of thinking about rewards: one may imagine that the relevant god just makes you pass, or one may imagine that the relevant god gives you some extra points -- whether the extra points are or are not sufficient to make the passing grade. I will be assuming the first way of thinking about rewards and punishments.

Then (57) should tell Joe that he would have passed even if he had not prayed -- that our god is not a god who rewards prayer, and also not a god who punishes non-prayer. Likewise, suppose that in the actual world, Joe does not pray and passes. Then (57) should tell Joe that he would have passed even if he had prayed -- that our god is not a god who rewards non-prayer, and also not a god who punishes prayer. Since Joe does not know (yet) which of these eventualities obtains, he rejects all four possibilities. Note that if this is the correct description of the import of (57), then (58), by parallel reasoning, rules out the exact same four possibilities -- so in fact Ivan's argument is redundant (I think this is the intuitively right result).

Now, let us come back to the main argument. Let's consider (57). First, make the following natural assumption: ' φ , whether or not ψ ' is true just in case 'If ψ , then φ , and if not ψ , then φ ' is true (I'll call this taking the whether-conditional *at face value*, see the next section for further discussion). If this is right, then

(57) If you pass, you'll pass whether you pray or not.

is equivalent to

(59) If you pass, then you will pass if you pray, and you will pass if you do not pray.

But it is natural to suppose that 'if φ , then ψ and θ ' is equivalent to 'if φ , then ψ , and if φ , then θ '. So, (11) entails

(60) If you pass, then you will pass if you pray, and if you pass, then you will pass if you don't pray.⁵⁸

58. By the way. Stalnaker's semantics gets (57) (and so (60)) right. So the situation is curiously symmetrical: Stalnaker gets the deviant readings right, and the normal readings wrong; all the other theories of conditionals get the normal readings right, and the deviant readings wrong.

But now, assuming Import-Export, (60) is equivalent to

(61) If you pass and pray, then you will pass, and if you pass and do not pray, then you will pass.

And (61) is a conjunction of two conditionals, in both of which the antecedent entails the consequent. So, assuming Import-Export, (61) is trivially true.

So here is the *argument against Import-Export*: on the face value assumption, Import-Export is inconsistent with our initial observation that (57) is informative in Pass/Fail. By parallel reasoning, Import-Export is also inconsistent with our initial observation that (58) is informative in Pass/Fail.

Furthermore, note that Heim's semantics is *also* inconsistent with our initial observations. So here is the *argument against Heim's semantics*: according Heim, (60) is logically true, since it is a conjunction of two instances of $\ulcorner \text{if } \varphi, \text{ then if } \theta, \text{ then } \varphi \urcorner$, and $\ulcorner \text{if } \varphi, \text{ then if } \theta, \text{ then } \varphi \urcorner$ is logically true on Heim's account.

The argument against both kinds of theories is really an argument epistemic context-shifting. If 'if' shifts the epistemic context, (60) must come out trivially true, contrary to fact. So the lesson of Pass/Fail is that nested whether-conditionals are sometimes *deviant*: the antecedent fails to shift the epistemic context.

The argument I just gave is negative: the assumption that 'if' always shifts the epistemic context leads to the wrong prediction. But I think we can also see that, in Pass/Fail, we evaluate Ivan's assertions *as if* the 'if' *does not change context*. So, suppose that in the actual world, Joe prays and passes. Question: is it relevant for the truth of (57) whether the actual world

contains a god who punishes those who do not pray? Answer: of course it's relevant! If there is such god, then (57) is *false*. But that means that some possibilities in which Joe does not pray and *fails* are relevant to the truth of (57), which is just to say that the consequent of (57),

(62) Joe will pass whether he prays or not.

is evaluated *not* in the epistemic context $c+Joe$ passes, but in a *wider* context, in a context which includes some Joe-fails-worlds. Although it is not the only possibility, it is *prima facie* plausible that in (57) 'if' does not shift the epistemic context at all.

To sum up: in or-to-ifs 'if' behaves normally, but in (some) nested whether-conditionals 'if' is deviant.⁵⁹ The *central problem* is to explain how normalcy and deviance can co-exist within one semantic theory. As we have just seen, no reasonable theory leaves room for the deviant cases. (And the only theory that gets the deviant cases right, Stalnaker's semantics, fails in the normal case) I will offer my solution in §7. But before I turn to that, I want to forestall a likely objection (§5), and then suggest an empirical generalization that will guide us to solution to the central problem (§6).

5 Taking whether-conditionals at Face Value

One might think: the argument against Import-Export and Heim works only if we take (57) and (58) at *face value*. But why do that? Let me explain why. First, here is my more precise, and

59. Why is epistemic context-shifting normal, and not shifting deviant, rather than the other way around? First, although I have no space to demonstrate this here, the normal behavior is indeed wide-spread (hence the intuitive appeal of Import-Export). Second, as we shall see, the deviant behavior arises only in very special circumstances (see §9 on singularity).

stronger, claim:

Face Value

' φ , whether or not Ψ ' is *strongly equivalent* to 'if Ψ , φ , and if not Ψ , φ '.

By 'strongly equivalent' I mean that the two sentences are true and false in the same contexts, are assertable and unassertable in the same contexts, and also share their presuppositions. (the last claim, presumably, follows from the second).

If *Face Value* were false, the semantics of whether-conditionals would still be an open question, but Pass/Fail would not be a direct counter-example to Import-Export, or to Heim's theory, because the argument works only once we 'translate' (57) and (58) using *Face Value* (that's the step from (57) to (11) above). Now, two arguments in support of *Face Value*, one direct, one indirect.

First, if *Face Value* were false, there should exist a context in which a whether-conditional is assertable, but the corresponding conjunction of indicatives is not (or vice versa). I cannot find such a context. That is a good reason to accept *Face Value*.

Second, the argument against Import-Export and Heim could as well be formulated without appeal to whether-conditionals. So, consider

(63) If you pass, you'll pass even if you don't pray. (uttered in Pass/Fail)

Initial observation: (63) is informative, just as (57) is, and, intuitively, for the same reason.

But (63) is predicted to be trivially true both by Import-Export and by Heim's theory. So we cannot get around the central problem by postulating a semantics for whether-conditionals

that diverges from *Face Value*, since the same problem arises with even if-conditionals.⁶⁰

6 Singularity

So far, we have precisely one example of deviance: (57) in Pass/Fail. To get a handle on the phenomenon of deviance, it would be very helpful to know under what circumstances it arises. Here, again, it is rather tempting to suppose that there is something special about whether-conditionals that is responsible for this. However, this temptation is to be resisted, as I will now show. So consider this nested whether-conditional, uttered in Pass/Fail:

(64) If Joe passes, he'll throw a party whether his neighbors like it or not.

I think it is clear that the truth of (64) does not depend on whether Joe will throw a party whether his neighbors like it or not in worlds in which he fails the exam. Supposing that his neighbors' sensitivity to loud music prevents him from throwing a party at one of those worlds does not stand in the way of (64)'s truth. Rather, the truth of (64) depends only on what happens in worlds in which Joe passes, and his neighbors are cooperative, and on what happens

60. I accept the following semantics for even if:

Even If

- i) 'P, even if Q' is true in w, c, just in case 'if Q, then P' is true in w, c.
- ii) 'P, even if Q', uttered at w, c, carries the presupposition that
 - 1) the context c is partitioned by some salient partition $R=\{R_1, R_2, \dots, R_n\}$, that is ordered by some salient relation $>$. (let's say $R_1 > R_2 > \dots > R_n$)
 - 2) $Q=R_i$, for some i. (some people think i must be n, but I disagree)
 - 3) 'If R_i , then P' is true in w, c, for all $R_i < Q$.

Example: If Joe touches alcohol, he'll get drunk even if he drinks one pint.

Let $R=\{\text{Joe drinks 5 pints, Joe drinks 4 pints, ..., Joe drinks one pint}\}$ (partitioning $c+\text{Joe touches alcohol}$). Intuitively, the conditional is assertable just in case all conditionals of the form 'if Joe touches alcohol, he'll get drunk if he drinks n pints' are true.

in worlds in which Joe passes, but his neighbors are not cooperative (note that I am in effect employing Import-Export here). (64) is normal, not deviant.

What lesson are we to draw from (64)? Clearly, the following generalization, by which we might have otherwise been tempted, is false:

False Generalization ('blame the 'whether'')

All and only nested whether-conditionals are deviant: the evaluation of the consequent proceeds in a non-shifted context.

The False Generalization, if true, could be the beginning of an explanation of deviance that appealed to some special feature of 'whether'. But (64) speaks against the False Generalization. I suggest, instead, that what distinguishes (57) from (64) is that in (57) the antecedent is identical with the nested consequent. Let's call nested whether-conditionals with this feature *singular*. Singular nested whether-conditionals are sentences of the form 'if φ , then φ , whether or not θ '. So I am proposing the following empirical generalization:

Singularity

All and only *singular* nested whether-conditionals are deviant: the evaluation of the consequent proceeds in a non-shifted context.

Let me stress that Singularity is a purely empirical generalization, and should be accepted (or rejected) by studying our actual use of nested whether-conditionals.⁶¹ I propose that we assume that it is singularity that is responsible for deviance, and try to explain how this might

61. Note that it is very natural to think that singularity extends also to those nested whether-conditionals 'if φ , then ψ , whether or not θ ' in which ' φ ' entails ' ψ '. I do not have the data to test this generalization.

come about.

7 The solution: Heim+

Let's take stock. Sometimes nested indicatives behave normally, and 'if' shifts the epistemic context -- as shown by (54) in the Election. Sometimes, nested indicatives behave deviantly, and 'if' somehow fails to shift context -- as shown by (57) in Pass/Fail. These two paradigms pull our semantic theory in opposite directions. That's the *central problem*. Finally, in the last section, I proposed the empirical generalization: that only singular whether-conditionals are deviant.⁶² If singularity is indeed the trigger of deviance, we would like an explanation of how this comes about.

I do not have a way to demonstrate that the following is the right solution. I think that one can fairly suspect it of being *ad hoc*. But the situation is so peculiar that any solution that actually explains the phenomena would be welcome. So here it is.

First, I propose that we, for the time being, adopt Heim's semantics.⁶³

Heim

An indicative 'if φ , then ψ ' is true in w, c , iff ' ψ ' is true in $\text{sim}(w, c+\varphi), c+\varphi$

Now recall that the deviance of (57) consisted in that

(57) If you pass, you'll pass whether you pray or not.

is evaluated in such a way that we evaluate 'you'll pass whether you pray or not' in a wider

62. Note, again, that I am just sticking with whether-conditionals for the purpose of argument. In fact, singularity extends beyond whether-conditionals -- witness even-ifs.

63. In §8, I will argue that in fact we *must* start with Heim's semantics.

context -- not in $c+Joe\ Passes$, but in c . Here, 'if' fails to shift the epistemic context.

Let's take singularity our guide: when the consequent of 'if φ , then ψ , whether or not θ ' is identical with the antecedent, i.e. when $\varphi=\psi$, we are forced to evaluate ' ψ , whether or not θ ' in c , rather than in $c+\varphi$.

Here is a natural hypothesis that seems like a step in the right direction: when we evaluate (57), we first, as we ought, update the context with the antecedent: we move from w,c to $\text{sim}(w, c+Joe\ passes)$, $c+Joe\ passes$. But now, suppose that indicative conditionals carry the following presupposition:

Open consequent presupposition

'if φ , then ψ ', evaluated in w, c , presupposes that c contains ψ -worlds and that c contains $\sim\psi$ -worlds.^{64,65}

With simple consequents, the presupposition is innocuous. It amounts to the harmless claim that, e.g.

(65) If Oswald did not kill Kennedy, Kennedy is dead.

is unassertable in our actual state of knowledge. (this is harmless because we would not want to assert (65) anyway).

But now return to evaluating (57). Once we have updated the context with 'Joe passes', the

64. The open consequent presupposition is used by von Stechow (1998, p. 8-9) to solve the Anderson problem. He calls it 'consequent variety', but suggests that it is weaker than a presupposition -- merely a 'presumption'.

65. Does 'if φ , then ψ ' presuppose that c contains both ψ -worlds and $\sim\psi$ -worlds? For my argument below to go through, it is necessary only that the existence of $\sim\psi$ -worlds is presupposed. So whether the presupposition is as strong as I make it is an open question.

open consequent presupposition is no longer satisfied by $c+Joe\ passes$, because that presupposition demands that the context contain worlds in which Joe passes, and worlds where he does not. I propose that we appeal here to *presupposition accommodation*. First, what is presupposition accommodation?

For example, suppose that our current context c is agnostic as to whether Joe is married, and also agnostic as to whether he has children. Now, someone says:

(66) If Joe is married, his children are at home.

The consequent carries the presupposition that Joe has children. Now, we have a choice. We can object to (44) with something like, 'Hey, we didn't know that Joe had children!', or we can take (44) to be informative, that is, we can accept the assertion of (44). And there is no doubt that we often do just that. Two questions: first, how can we accept (44), given that the presupposition is not satisfied by the context of utterance? Second: what is it that we learn when we accept (44), or put differently, how do we update the context?

The answer to the first question is pretty uninformative: the answer is that we need to posit a new mechanism, *presupposition accommodation*. The mechanism works as follows: when φ is uttered in a context c , and φ carries a presupposition that ψ , and the context does not satisfy ψ (that is, some worlds in c are ψ -worlds, and some not), we pragmatically restrict the context to $c+\psi$, and evaluate φ in the new, restricted context. The result of the update is $c+\psi+\varphi$. As I understand it, the mechanism is necessarily pragmatic: that is, it is still the case that (44) does not have a semantic value as uttered in c , it's just that we voluntarily choose to restrict the context, so as to allow it to have a semantic value.

There is one more important detail that (44) brings out. When we accept the assertion of (44) in the context that is agnostic, as I said, about the existence of Joe's children, we in fact have two ways in which we could update.

We can accommodate *globally*. We can first update the context with 'Joe has children'. The result of this will be the following update: $c + \text{Joe has children} + \text{If Joe is married, his children are at home}$. The result is a context in which it is certain that Joe has children.

We can also accommodate *locally*. We can first shift to the context $c + \text{Joe is married}$ (beginning to interpret the conditional), and then accommodate. The result is a context in which it is not certain that Joe has children, but certain that if Joe is married, he has children. (one can say: it is as if we accommodated globally, but not the presupposition of the consequent, but a different, conditional proposition: if Joe is married, he has children).

Now let's come back to (57). First, observe that in the case of (57), we cannot accommodate the open consequent presupposition globally. Were we to do it, our labor would immediately have been lost, since, once we update by the antecedent of (57), the context would again be one that does not satisfy the open consequent presupposition. So we must accommodate locally.

But things are worse still. The familiar cases of accommodation are cases where the context (global or local) is *restricted* -- updated by some proposition (as with (44)). But in our case, the open consequent presupposition cannot be accommodated by *restricting* the (local) context $c + \text{Joe passes}$ -- our problem is that the context is already too restricted. I suggest that in this case, we accommodate *up*, instead of *down*: we accommodate not by restricting the context,

but by expanding it.⁶⁶

Up-accommodation of the open consequent presupposition

If 'if φ , then ψ ' is evaluated in context w , c , and c entails ψ , the context is, through local accommodation, expanded to $c-\psi \equiv c \setminus \{\psi\}$.⁶⁷

It is important to point out that, while it suffices for the cases we are concerned with, the formulation just given may need to be revised. So, for example, consider the possibility that $c-\{\psi\}$ entails ψ . The above definition suggests that in such a case we still do not succeed in accommodating the open consequent presupposition, and so presumably irremediable presupposition failure would result. But it may well be the case that in such cases we in fact successfully accommodate. If so, the up-accommodation mechanism would have to be complicated further.⁶⁸

To sum up my proposal so far: I am proposing that we enrich Heim's theory. Accepting open consequent presupposition and up-accommodation gives us the following theory, call it Heim+:

Heim+

i) *Truth-conditions*

66. Do we have examples of accommodation by expansion? Perhaps we do. When one believes that P , and someone else asserts that Q , such that Q presupposes $\sim P$, one can only accommodate by expansion, if one accommodates at all. (thanks to Harold Hodes for pointing this out).

67. Note that here it is crucial that c is a set of propositions, rather than a set of worlds. If we removed all the Q -worlds from c , we would be left with the empty set, in cases where up-accommodation is triggered by the failure of the open consequent presupposition.

68. Note also that if we wanted up-accommodation to guarantee that the resulting context satisfies the open consequent presupposition, there would of course be, in principle, many ways to do it: if a set of propositions Γ entails φ , there can, in principle, be many ways of subtracting some propositions from Γ in such a way that $\Gamma \setminus \{\psi_1, \psi_2, \dots\}$ no longer entails φ . Clearly there is more work to be done on exactly how up-accommodation works.

An indicative 'if ϕ , then ψ ' is true in w, c , iff ψ is true in $\text{sim}(w, c+\phi), c+\phi$

ii) *Presuppositions*

An indicative 'if ϕ , then ψ ', evaluated in w, c , has the presupposition that ψ is open in c , i.e. that $c \models \psi$ and $c \models \sim\psi$.

iii) *Presupposition Accommodation*

When the open consequent presupposition is not satisfied by the local context, the local context accommodates up: $c \Rightarrow c-\psi$.

Heim+ is just like Heim in non-singular cases (e.g. (64) and (54)), and so Heim+ gives us the right results in the Election, and for or-to-if conditionals generally. But Heim+ also gives us the right result for Pass/Fail: (57) and (58) trigger the presupposition accommodation mechanism, and their consequents are evaluated in expanded local contexts. This resolves the Central Problem.

To illustrate, here is how Heim+ handles (57). Let's evaluate (57), step by step, in Pass/Fail.

(57) If you pass, you'll pass whether you pray or not.

i) we update the context w, c , to $\text{sim}(w, c+\text{Joe passes}), c+\text{Joe passes}$.

ii) we encounter the open consequent presupposition failure, trying to evaluate 'Joe will pass, whether he prays or not', in $\text{sim}(w, c+\text{Joe passes}), c+\text{Joe passes}$. So we accommodate up to $\text{sim}(w, c+\text{Joe passes}), c$.

iii) we evaluate 'Joe will pass, whether he prays or not' in $\text{sim}(w, c+\text{Joe passes}), c$. Pass/Fail is god-agnostic. So suppose first that $\text{sim}(w, c+\text{Joe passes})$ is a world with a god who punishes non-prayer: if you don't pray, he makes you fail. Then 'Joe will pass, whether

he prays or not' will be FALSE in $\text{sim}(w, c + \text{Joe passes})$, c . On the other hand, suppose that $\text{sim}(w, c + \text{Joe passes})$ is a world without a god, then 'Joe will pass, whether he prays or not' will be true in $\text{sim}(w, c + \text{Joe passes})$, c .

So, (57) is informative: it tells Joe that god does not punish non-prayer (and similarly for other possibilities -- rewarding prayer, etc.).

8 Why Heim?

Now we have an explanation of deviance: singularity triggers a mechanism of presupposition accommodation *via* the open consequent presupposition. We have just seen how the mechanism, combined with Heim's semantics, handles Pass/Fail. But the mechanism itself is theory-neutral, and can be combined with any semantic theory of indicative conditionals. So the next natural question is: can our success with Heim+ be replicated by fixing up other theories with the proposed mechanism? I think that the answer is NO, but I do not have a general argument that would rule out success in every case. Instead, I will illustrate two cases of failure -- I'll show that McGee+ and Kratzer+ fail to make the right predictions in Pass/Fail.

McGee's theory goes as follows.

McGee

An indicative 'if ϕ , then ψ ' is true in w , c , iff ' ψ ' is true in w under the set of hy-

potheses $c+\varphi$.⁶⁹

(The recursive definition of 'true under the set of hypotheses Γ ' is in the footnote. But the basic idea is this: 'An atomic sentence is true in w under the set of hypotheses Γ iff it is true in the possible world most similar to w in which all the members of Γ are true... Finally, $\neg\varphi\rightarrow\psi$ is true in w under the set of hypotheses Γ iff ψ is true in w under the set of hypotheses $\Gamma\cup\{\varphi\}$.' McGee 1985)

McGee's theory looks superficially like Heim's, but there is a core difference, that comes out most clearly if we trace what McGee says about right-nested indicatives. Take an arbitrary right-nested indicative

(67) \neg if φ , then if ψ , then θ .

McGee says that (42) is true in w , c just in case \neg if ψ , then θ is true in w on the hypothesis that $c+\varphi$, which, in turn, is true just in case $\neg\theta$ is true in w on the hypothesis that $c+\varphi+\psi$ -- that is, just in case the closest $c+\varphi+\psi$ -world to w is a θ -world. Contrast this with Heim's prediction that (42) is true in w , c just in case $\neg\theta$ is true in $\text{sim}(\text{sim}(w, c+\varphi), c+\varphi+\psi)$, $c+\varphi+\psi$, Heim and McGee both shift the epistemic context, but Heim (like Stalnaker) also shifts the

69. Here is how McGee explains it: 'It is not hard to modify the Stalnaker semantics so that it has the right logical features. Instead of a simple notion of truth in a world, we develop a notion of truth in a world under a set of hypotheses. To be simply true in a world is to be true in that world under the empty set of hypotheses. If there is no world accessible from w in which all the members of Γ are true, then every sentence is true in w under the set of hypotheses Γ . Otherwise we have the following: An atomic sentence is true in w under the set of hypotheses Γ iff it is true in the possible world most similar to w in which all the members of Γ are true. A conjunction is true in a world under a given set of hypotheses iff each of its conjuncts is. A disjunction is true in a world under a set of hypotheses iff one or both disjuncts are. $\neg\varphi$ is true in w under the set of hypotheses Γ iff φ is not true in w under that set of hypotheses. Finally, $\neg\varphi\rightarrow\psi$ is true in w under the set of hypotheses Γ iff ψ is true in w under the set of hypotheses $\Gamma\cup\{\varphi\}$. Thus to evaluate whether $\neg\varphi\rightarrow(\psi\rightarrow\theta)$ is true under the set of hypotheses Γ , we add first φ and then ψ to our set of hypotheses, and we see whether θ is true under the augmented set of hypotheses $\Gamma\cup\{\varphi, \psi\}$.' (McGee 1985, p. 469)

world of evaluation. (McGee's 'if' also shifts the world of evaluation -- $\neg\theta$ is evaluated in the closest $c+\phi+\psi$ -world, but McGee shifts the world of evaluation only once, at the end of a chain of 'if's). It is easy to see that this allows McGee to validate Import-Export, which was indeed his goal in devising the semantics.

Now, we can modify McGee's semantics just as we did Heim's. The result will be McGee+. What predictions does McGee+ make in Pass/Fail? McGee+ gives the wrong result. Here is the argument. Consider Pass/Fail**, which is just Pass/Fail together with the knowledge that there are no gods. Joe ought to think both (57) and (58) true, no matter which world in c is actual. That's because Pass/Fail** has no gods that listen to prayers. But not according to McGee+. Here is the idea. Recall:

(57) If you pass, you'll pass whether you pray or not.

Now, McGee+ suggests we evaluate (57) as follows:

- i (5) is true in w under the set of hypotheses c just in case 'you'll pass whether or not you pray' is true in w under the set of hypotheses c +Joe passes.
- ii We encounter open consequent presupposition failure, and shift from the set of hypotheses c +Joe passes to the set of hypotheses c .
- iii We evaluate 'you'll pass whether or not you pray' in w under the set of hypotheses c .

Now, if w is a world in which Joe does not pass (certainly a possibility!), then 'Joe will pass, whether he prays or not' will be false in w , c . That's the wrong result. What is responsible for failure? I think the answer is pretty clear: the problem is that when the presupposition accommodation mechanism is triggered, McGee+ reverts to the original context of evaluation w , c .

Heim+, by contrast, reverts to $\text{sim}(w, c+\text{Joe passes}), c$. In other words, what allows Heim+ to get things right is that Heim+ shifts the world of evaluation. McGee+ does not shift the world of evaluation, and so gives the wrong result.

Similar remarks apply to strict-conditional approaches, like Kratzer's theory. Very briefly:

Kratzer-like semantics

'if φ , then ψ ' is true in w, c just in case ' ψ ' is true in $w, c+\varphi$ for every w in $c+\varphi$.

Again, let's consider Pass/Fail**, and go through the reasoning:

- i (5) is true in w, c just in case 'you'll pass whether or not you pray' is true in $w, c+\text{Joe passes}$ for every w in $c+\varphi$.
- ii We encounter open consequent presupposition failure, and shift from the context $c+\text{Joe passes}$ to the context c .
- iii We evaluate 'you'll pass whether or not you pray' in w, c , for every world in c .

So 'Joe will pass, whether he prays or not' will be false in w, c , because c contains worlds in which Joe does not pass.

To sum up. The argument in favor of Heim+ is of course far from complete. But the failure McGee+ and Kratzer+ points to the crucial role that the shifting of the world-parameter plays in Heim+. If the shifting of the world-parameter is indeed essential, then Heim+ is the right theory.

9 Conclusion

Now we have a theory, Heim+, that accounts for all the data. It needs two new principles: open consequent presupposition, and up-accommodation. Of these, up-accommodation, if true, is the most extraordinary. Since up-accommodation (if it exists) is not well-understood, it is hard to claim that I have a knock-down argument for Heim+. It would be wiser to take the main result dialectically: as an incentive to put up-accommodation on a firmer footing, or else to find a new theory that accounts for singularity but does not appeal to up-accommodation.

Chapter 4

The Unified Theory

1 Introduction

A single word, *if*, is used to express conditional thoughts in English. One should therefore expect research on the conditional to produce a unified theory of *if*. So it is *prima facie* surprising that most of the research into the conditional construction has broken into two streams, one pursuing the semantics of the indicative conditional, the other of the subjunctive. To be sure, interest in indicatives goes along with interest subjunctives -- after all, they are similar in many ways -- but attempts at a unified analysis are few. Not only are such attempts few, but, what is more, the investigation of conditionals is not, for the most part, guided by the aim of (ultimately) building a unified theory. Why is that?

Certainly the reason is not that a lexical ambiguity theory is thought plausible: it has, so far as I know, no adherents.⁷⁰ Rather, the current practice of concentrating on one or the other kind of conditional follows the simple thought that, given that indicatives and subjunctives seem importantly different, they are best investigated separately (I will discuss the motivation for this view in section 2).

70. Although with a view like Lewis' it is hard to see how to avoid commitment to outright lexical ambiguity.

This two-pronged methodology is accompanied by the thought that, given the differences between indicatives and subjunctives, the most one can hope for in the way of unifying the semantics of *if* is some sort of perspicuous *similarity* in the accounts of indicatives and subjunctives. And, one might think, we have already discovered all the similarity there is, and so the quest for a unified semantics has ended with a half-victory and a half-defeat. (As we shall see, Kratzer's account is like that.)⁷¹

The traditional two-pronged approach has produced an enormous amount of interesting results; there is nothing wrong with investigating indicatives *in vacuo*, or subjunctives *in vacuo*. But the fact that, at the end of the day, *if* must have a unified semantics is itself a piece of evidence that ought to be used in our work on the conditional. So in this paper I pursue the opposite methodology: I will ask what we can learn about the semantics of indicative and subjunctive conditionals from the fact that *if* must have a unified semantics. No currently available theory accommodates this demand, and so we need a new theory.

To pursue this line of thought one needs to have an idea of what it is for a semantics of *if* to *be* unified. I think two demands on unity have been recognized. One is the vague demand that we need a single lexical entry for *if*. The demand is vague because it fails to produce real constraints if the lexical entry itself is allowed to be disjunctive in some way. The other demand on unity is the need to explain indicative-to-subjunctive inference: the inference from 'if it rains, I will stay at home' to a later utterance of 'had it rained, I would have stayed at

71. The talk of 'kinds' of conditionals here is loose. Perhaps there are more than two kinds, and perhaps the right division is not into indicative and subjunctive, but into biscuit, deontic, circumstantial, subjunctive, etc. I will ignore these issues here. For the purposes of this paper, the focus is on epistemic indicatives and the subjunctives (the latter taken by everyone to form a unified class).

home'. While it is not clear that we have a satisfying explanation of the indicative-to-subjunctive inference, I shall not pursue this question here.

The starting point of this paper is that there is another job for a unified semantics of *if*: to explain the pattern of presuppositions of indicative and subjunctive conditionals. Demanding that our unified theory of the conditional does this job has important consequences for the shape that the unified theory can take.

Before articulating the new demand, I want to review the main motivation for the traditional two-pronged approach: the familiar Kennedy minimal pairs:

(68) If Oswald did not kill Kennedy, someone else did.

(69) Had Oswald not killed Kennedy, someone else would have.

My main claim is that the traditional understanding of the lesson to be drawn from the Kennedy pairs encourages two connected thoughts, both, I will ultimately argue, mistaken: one, that the best one can do by way of providing a unified semantics of *if* is a story according to which indicatives and subjunctives are very similar; two, that indicatives and subjunctives are fundamentally different in that indicatives deal in epistemic modality, while subjunctives deal in metaphysical modality (call this intuition the *modal gap* intuition). I will be occupied with the question of unity throughout sections 3-5, and pick up the question of just how many modalities are involved in sections 6-7. In the end, I will suggest that a variant of Stalnaker's semantics does best as a unified theory, and that both indicatives and subjunctives, in an important sense, give us a view of the same modal realm.

2 Kennedy pairs

The first basic fact one learns about conditionals is that there are (at least) two kinds of conditionals: indicatives and subjunctives. So, consider (Adams 1970):

(68) If Oswald did not kill Kennedy, someone else did.

(69) Had Oswald not killed Kennedy, someone else would have.

Call (68) and (69) a *Kennedy pair* (it is convenient to generalize, and call any similarly related indicative-subjunctive pair a Kennedy pair). What are our *first* intuitions about this pair, given our current historical knowledge? The most natural thought is that (68) is surely true, because we know that Kennedy was shot, but that (69) is likely false, because we think it is likely that Oswald acted alone.

Focus on the Kennedy pair, and the acceptance of the first judgments on (68) and (69), have had two profound effects on our thinking about conditionals.

The first thought goes as follows. The Kennedy conditionals, since they (likely) differ in truth-value, have *different truth-conditions*.⁷² In particular, indicative conditionals are *epistemic* in flavor, while subjunctive conditionals *metaphysical* in flavor (what we know is relevant to the truth-value of (68), and not of (69)). It is commonly thought that the indicative reports on the current epistemic situation, while the subjunctive makes a claim that is independent of what

72. Note that my discussion in this section neglects the question of distinctive subjunctive morphology. The difference between (68) and (69), one may object, should be attributed to the peculiar morphology that appears in (69), not to the differing truth-conditions of the indicative *if* and the subjunctive *if*. The objection is right, and I will turn to the role of morphology in section 4. However, the bulk of the early literature on conditionals also neglected the role of morphology, and so my discussion here is, in this sense, historically accurate. In this section, I am pursuing a genealogical question: how did we come to think that there is a modal gap between indicatives and subjunctives?

the current epistemic situation is.⁷³ Indicatives are context-sensitive, while the subjunctives are not. Indicative and subjunctive conditionals afford us access to two different modal realms, that of epistemic and metaphysical possibility.⁷⁴ So, for example, in some circles, it is common to speak of taking the antecedent *as actual*, or *as counterfactual*, these two phrases being cues for interpreting the conditional as relating to metaphysical modality (second-intension), or epistemic modality (first-intension).⁷⁵ The intuition is that there is a *modal gap* between indicatives and subjunctives. It is not necessitated by the Kennedy pairs, but strongly encouraged by these and similar examples of indicative/subjunctive divergence.

The second thought naturally follows the first. Given that indicatives and subjunctives report on two different modalities, the semantics of *if* cannot be *too* unified: the most one can expect is a high degree of structural similarity. Just what kind of similarity is involved is a further question. Kratzer's influential view is that the conditional is context-sensitive -- in some contexts, it reports on the epistemic modality, and in other contexts, on the metaphysical modality (see section 4).

Kennedy pairs are paradigmatic and suggestive -- suggestive of the existence of a modal gap, and suggestive of only a weak unity in the semantics of *if*. That is why it is important to see that the natural first reaction to the Kennedy pairs above, as described, is highly misleading. Seeing that it is misleading will set up the central question of this paper.

73. Gibbard 1981 is the *locus classicus* for the epistemic dependence of indicatives.

74. E.g. von Fintel 1998, p. 2: 'Most researchers will however agree in some form or other that indicative and subjunctive conditionals differ in *at least* the epistemic status of their domain of quantification.'

75. See Chalmers 2006.

3 Presuppositions

Consider the Kennedy pair again:

(68) If Oswald did not kill Kennedy, someone else did.

(69) Had Oswald not killed Kennedy, someone else would have.

The natural reaction, as we saw above, is to think that (68) is true, while (69) is likely false (probably Oswald acted alone). But this way of expressing our attitude toward (68) and (69) hides the fact that we are never in a position to gauge our judgment on both members of a Kennedy pair *in the same context*. So, to evaluate (69), we must take it for granted that Oswald killed Kennedy. But we cannot utter (68) in that same context -- (68) requires that it be open/possible/not taken for granted that Oswald killed Kennedy. Our first intuitions about the Kennedy pairs, then, rest on a suspicious *slide*: passing from the indicative to the subjunctive, we silently adjust the context in which we evaluate them, and only in this way can we get the natural intuition of truth-conditional divergence. (the line between contexts in which (68) is felicitous, and the contexts in which (69) is felicitous is quite thin: doubt that Oswald is the culprit is enough to make one switch from using (69) to using (68) -- but the line is there).

In fact, our illicit slide when evaluating the Kennedy pair is an instance a general phenomenon: arguably, indicatives and subjunctives find themselves in a *complementary distribution*.⁷⁶

76. Everyone agrees, I think, that indicatives and subjunctives are in a complementary distribution. This leaves open the question just what that distribution is: in particular, it leaves open the question whether all subjunctives are counterfactual. On this issue, see von Stechow 1998.

Just for the record: I hold the minority view that all subjunctives are counterfactual.

That is, the contexts in which indicatives are felicitous, subjunctives are not, and vice versa.⁷⁷ I will take this as an empirical given.

It is natural to attribute the distribution of felicity/infelicity judgments to the presence of *presuppositions*.⁷⁸ We can say, then, that indicatives presuppose that the antecedent is (epistemically) possible, whereas subjunctives presuppose that the antecedent is not (epistemically) possible.^{79,80} While much remains to be said about the presuppositions of conditionals, the mere fact that the distribution of indicatives and subjunctives is complementary throws doubt on the two thoughts prompted by our first reaction to the Kennedy pair, described in section 2.

So, first of all: if there were a context in which both (68) and (69) were felicitous and differed in truth-value this would be good evidence that indicatives and subjunctives have different truth-conditions. But there is no such context, and so this claim is suspect. Since the intuition that indicatives and subjunctives report on two different modal realms is based on the supposed divergence in the truth-conditions, that intuition is thus suspect as well. To be sure, the modal gap intuition may well be more entrenched than that, and have other sources besides the Kennedy pairs. I will argue against the modal gap claim directly in section 7. (A

77. Are there contexts in which neither indicatives nor subjunctives are felicitous? Presumably, yes; but I want to bracket that question, and concentrate on the competition between indicatives and subjunctives.

78. Stalnaker 1975 thinks it is an implicature, not a presupposition. I think Stalnaker's reasons are not convincing. I will assume, with von Stechow 1998, that it is a presupposition.

79. Perhaps indicatives and subjunctives presuppose more than that. Iatridou 2000 thinks that a subjunctive presupposes that the consequent is false.

80. This claim hides a significant assumption. The well-known counter-example to the claim that subjunctives are counterfactual is due to Anderson: 'If Jones had taken arsenic, he would have shown just the symptoms that he does in fact show' (uttered in a context in which it is not known whether Jones took arsenic). My hypothesis is that the Anderson example does not in fact show that some subjunctives are not counterfactual: a different explanation is available. But I want to bracket this complicated issue here.

vaguer intuition remains: that we somehow do different things when we go about evaluating (68) and (69) -- that much I happily grant).

Second: since the thought that indicatives and subjunctives have different truth-conditions is suspect, so is the thought that the semantics of indicatives and subjunctives is unified only by some sort of similarity.

The main motivation for the traditional approach rests on a slide, and is suspect. Let me stress that I do not mean this to be a self-standing criticism of the traditional approach -- only a motivation to ask further questions. The traditional lesson of the Kennedy pairs ignores the facts about the presuppositions of conditionals. So the natural question to ask is: why do indicatives presuppose that the antecedent is open (epistemically possible) and subjunctives that the antecedent is not open (epistemically impossible)? I want to show that pursuing this question will yield an important constraint on any theory of the conditional.

The simplest idea is that the semantic theory of indicatives and subjunctives can somehow explain their presuppositions.⁸¹ And it does seem that in the case of indicatives this is a workable proposal.

Although no single answer is accepted on all sides, the answer to the first question, why indicatives presuppose that the antecedent is possible, is intuitively clear. On any currently available semantic theory of the indicative, indicatives with epistemically impossible antecedents are somehow defective. For some theories, for example the probabilistic Adams-inspired view, such indicatives are undefined (because the conditional probability of the conse-

81. The paradigm case is the Strawson-Frege theory of the definite article. The semantics of *the* explains why it is infelicitous in cases where there is no unique satisfier of the definite description.

quent on the antecedent is undefined). On other views, e.g. Kratzer's, they are trivially true. But trivial truth in the latter case is just a by-product of the semantics. There is no plausible justification for counting indicatives with epistemically impossible antecedents among logical truths. So our understanding of the semantics of the indicative suggests that indicatives with epistemically impossible antecedents are defective. It is plausible that in this case we are dealing with presupposition failure. So there remains a question about the formal mechanism.

Here one approach is to explain the presuppositions of indicatives by appeal to the semantics. (the alternative approach, that appeals to the special morphology present in conditional sentences will be taken up in section 5). One possibility is that the semantics of an indicative conditional is akin to the semantics of definite descriptions: an indicative presupposes that there is a unique closest world in the same way a definite description presupposes that there is a unique salient satisfier of the description (this is the proposal of Schlenker 2004). Another possibility is that the indicative conditional is like a universal quantifier over the domain of the epistemically possible worlds, and carries an existence presupposition just like the quantifier presupposition that the domain is not empty ('All seniors in the class must write a term paper' presupposes that there are seniors in the class) (see von Stechow 1998).

So, really, the interesting question is: why do subjunctives carry the presupposition that they do, *viz.* that the antecedent is not actual? Here the situation is very different from the indicative case: no semantic theory of the subjunctive, by itself, helps explain how subjunctive presuppositions arise. Just as an illustration, consider the Lewis-Stalnaker semantics for subjunctives: a subjunctive 'if P, then Q' is true at w just in case the closest P-world to w is a Q-

world. But, given an absolute (as opposed to context-sensitive) similarity ranking on worlds, the Lewis-Stalnaker semantics should be happy producing truth-values for subjunctives even in conversational contexts that *are* compatible with the truth of the antecedent. The limitation to counterfactual contexts appears entirely arbitrary.

No semantics for subjunctives taken as a self-standing theory can explain the presuppositions of subjunctives (this is a generalization that of course I have not defended -- I see no way to defend it, other than by going through the available proposals). What can then explain these presuppositions? One direction in which one might look is unified theories of the conditional. The intuitive motivation is this: since the presuppositions of indicatives and subjunctives are complementary, one can expect the explanation of this fact from a theory that deals with both kinds of conditionals within a single framework. The most widely accepted unified theory is Kratzer's.

4 Kratzer's unified theory

There is in fact a theory that claims to be a unified account of the conditional, and we should examine it first -- it is Kratzer's view. First, consider

(70) John may remain seated when the Queen walks in.

(70) may express the thought that John, for all we know, may remain seated when the Queen walks in. This is an *epistemic* reading of (70). But (70), in a different context, may express the

thought that John is *permitted* to stay seated when the Queen walks in. In this case, (70) has a *deontic* reading. In both cases, we may think, *may* is an existential quantifier over worlds: in the first case, over worlds consistent with what is taken for granted in the context, in the second case, over worlds consistent with ought-facts holding in the context. In both cases, *may* quantifies, clearly, over the set of worlds *salient* in the context of utterance. If the hearer of (70) is not sure which context is most salient, she may well ask for a clarification: 'do you mean that John is *allowed* to remain seated, or that he *may, for all you know*, remain seated?' This is a clear case of contextual ambiguity: the epistemic context is relevant in the first case, a deontic context in the second case.

According to an influential proposal by Angelika Kratzer, indicatives and subjunctives are related in just the way the two readings of *may* above. Kratzer proposes the following truth-conditions:

(71) $\llbracket \text{if } P, \text{ then } Q \rrbracket^{f, g, w} = 1$ iff all the $f(w)$ -minimal P -worlds in $g(w)$ are Q -worlds.

Here, $f(w)$ and $g(w)$ are two contextual parameters -- the ordering source, and the modal base. Both $f(w)$ and $g(w)$ are thought of as sets of propositions, so we should think of the worlds *in* $g(w)$ as the set of worlds in which all the propositions in $g(w)$ are true. According to Kratzer, when we are dealing with an indicative, $f(w)$ is empty, and $g(w)$ is epistemic: the result is that all the worlds in $g(w)$ are f -minimal. So the truth-conditions in (71) amount to: an indicative is true iff all the P -worlds consistent with what is taken for granted in the context of utterance are Q -worlds. When dealing with a subjunctive, on the other hand, $g(w)$ is empty (so all the worlds in W are in $g(w)$), and $f(w)$ encodes the Lewis-Stalnaker similarity metric. So, a sub-

junctive is true iff all the closest P-worlds out of all the worlds in W, the universe of possible worlds, are Q-worlds.

Kratzer's proposal amounts to a contextualist answer to the unity question: indicatives and subjunctives have a unified semantics in the sense that they share a single logical form, and the different truth-conditions result from different values of the contextually supplied modal base g and ordering source f .

Now, let's come back to our main question: why are indicatives and subjunctives in complementary distribution? We saw that this question really reduces to: why do subjunctives have the counterfactuality presupposition?

Kratzer, so far as I know, does not explicitly discuss this question, so the following is my reconstruction that is inspired by the analogy, in Kratzer's framework, between indicatives and subjunctives on the one hand, and the epistemic and deontic readings of *may* on the other. I suggest that the only answer Kratzer can offer within her contextualist framework is this: *metaphysical f and g are conversationally salient when the antecedent is inconsistent with the common ground, and the epistemic f and g are salient when the antecedent is consistent with the common ground.* It is the only answer available because the choice of f and g is the only thing that distinguishes indicatives and subjunctives, and f and g are external parameters, supplied by the context.

For all that's been said this answer is *ad hoc*. Why, we may ask, cannot a metaphysical modal base and ordering source be salient in a conversation that does not take the falsehood of the antecedent for granted? In the abstract: that, in a certain context w , C , all the P-worlds

in C are Q-worlds, and that all the closest P-worlds to w are Q-worlds are two *independent* pieces of information: so why can't one be interested in finding out both?

Here is a concrete example. Consider Gibbard's riverboat scenario:

Sly Pete. Sly Pete and Mr. Stone are playing poker on a Mississippi riverboat. It is now up to Pete to call or fold. My henchman Zack sees Stone's hand, which is quite good, and signals its content to Pete. My henchman Jack sees both hands, and sees that Pete's hand is rather low, so that Stone's is the winning hand. At this point, the room is cleared. [...] Zack knows that Pete knew Stone's hand. He can thus appropriately assert "If Pete called, he won." Jack knows that Pete held the losing hand, and thus can appropriately assert "If Pete called, he lost." (Gibbard (1981), p. 231).

Naturally, Pete folds. Now, concentrate on Zack's conditional. It is true, we may suppose, despite the fact that the nearest world in which Pete calls, he loses. And when the news that Pete folded becomes available, Zack may truthfully say: had Pete called, he would have lost. But note also that Zack may very well be interested in the truth of the Kratzer-subjunctive even *before* he hears the news of the outcome of the game. So, in the Sly Pete scenario, one might be interested *both* in the question whether all the worlds in the common ground in which Pete plays are Pete-wins-worlds, *and* in the question whether the closest Pete-plays world is a Pete-wins-world.⁸² Given the semantics of the Kratzer conditional, the prohibition on simultaneous felicity of the indicative and the subjunctive is unmotivated. So Kratzer's theory does not provide a good explanation of the presupposition pattern of indicatives and

82. This is a particularly interesting question in the riverboat scenario, since one might have the intuition that 'if Pete plays, he will win' is assertible, while the corresponding 'Had Pete played, he would have won' is not -- i.e. the scenario seems to be a counterexample to the indicative-subjunctive inference (and I think it is!).

subjunctives.^{83,84}

The preceding argument, one might think, is not enough. Consider the case of *may*. True, 'it may be that P', in a single context, expresses either the deontic or the epistemic reading (or perhaps some other). But one can, nevertheless, be interested in both. So, for example, one can, in a given context, be interested in whether John may sit in *both* the deontic and epistemic sense. In such cases, I can make my interest clear: I can say that I would like to know whether John is permitted to sit, and whether he might sit. So it is possible to be interested in two readings, but be unable to express them without a declaration to the effect that the question under discussion should change (say, from epistemic to deontic). In such cases, the fact that the context makes one or the other reading expressible, but not both, is the result of the poverty of one's language -- one could easily imagine a language in which the deontic and epistemic readings of *may* were expressed with different linguistic devices.

Poverty of language is the only thing that stands in the way of expressing, in a single context, both the deontic and the epistemic reading of *may*. Is the case of the indicative and the subjunctive similar? Not at all. Nothing like this is available in the case of the conditional. The linguistic means of expressing the indicative and the subjunctive are already different, due to

83. Further, it might very well be that there are pairs of contexts that are epistemically indistinguishable (=share the same epistemic context), but such that in one, the indicative reading is relevant, and in the other, the subjunctive reading is relevant. In such a case, the twin-presupposition view will not allow the felicitous utterance of the indicative in the first context, or the felicitous utterance of the subjunctive in the other. But WHY? What one wants to say here is: the subjunctive reading expresses a perfectly coherent proposition, and so does the indicative reading; if the twin-presupposition view prohibits one of them, the prohibition is *arbitrary*. (the arbitrariness is most pressing in cases where the indicative is felicitous, but the subjunctive is not; when the indicative is infelicitous, on the twin-presupposition view, we do not lose much: only the opportunity to assert a trivial truth, or a trivial falsehood).

84. The conclusions of this section are equally applicable to Gillies 2007, since it, too, deals in two contexts: the epistemic and the metaphysical, and so, again, only conversational relevance determines the choice of context.

the presence of subjunctive morphology, and yet Zack cannot say that he would like to know whether Pete will win if he calls, *and also* whether he would have won if he called. This twin interest just seems incoherent. And yet Kratzer's semantics fails to explain this.

5 Iatridou: explaining the presuppositions

It is not surprising, one might think, that Kratzer's theory does not provide the answer to the question we are asking. After all, indicatives and subjunctives differ in the mood, tense, and aspect morphology (depending on the language). But Kratzer's semantics abstracts from the role these elements might have in contributing to the semantic value, or to the presuppositions of a conditional construction. It is unfair, one might object, to ask Kratzer's semantics to do a job it was not designed to do. So we must consider how to pursue our guiding question of explaining the presuppositions of indicatives and subjunctives when account is taken of these additional elements. There are several proposals in the literature. I will consider Iatridou 2000 (cf. also Schlenker 2005).

Iatridou claims that 'CF [counterfactual] conditionals differ from non-CF conditionals in meaning as well as in form. I will argue that it is possible to attribute the difference in meaning to a systematic difference in verbal morphology' (Iatridou 2000, p. 232). As Iatridou convincingly argues, subjunctive conditionals in English and Modern Greek carry a layer of

fake past and a layer of fake imperfective aspect.⁸⁵ *Fake past* is a past tense that does not shift the time of evaluation of the clause to the past (and likewise for the fake imperfective aspect). So,

(72) If you left tomorrow, you would get there the next week. (Iatridou 2000, p. 235)

entertains a future possibility -- the past tense of the antecedent does not receive its normal temporal interpretation. Iatridou's proposal is that '... the past tense morpheme always has the same meaning, but the domain it operates on varies according to the environment.' (p. 245). So, in normal contexts, the morpheme conveys the presupposition of pastness, but in certain special contexts, it conveys a different presupposition, viz. the presupposition of counterfactuality. The details do not matter in the present context. For our purposes, Iatridou's proposal is that it is reasonable to interpret the past tense morpheme ('or rather, the feature whose phonetic realization we call the "past tense morpheme"', p. 246) as carrying the presupposition that the world(s) in which the antecedent is being evaluated are outside the common ground.⁸⁶

Note that Iatridou's proposal, although superficially of the same basic form as Kratzer's (the context disambiguates which reading is relevant: epistemic or metaphysical, past or counterfactual), escapes the criticism of the previous section. In Kratzer's case, we are asking the

85. Iatridou, following Stalnaker, thinks that what I have been calling the presuppositions of subjunctives are implicatures. But she does not give convincing reasons for this proposal. I will ignore this divergence.

86. Iatridou's proposal might be thought to carry a significant cost: it assumes that the past morphology expresses pastness by carrying the presupposition that the reference time of the clause is in the past. Such an approach has to compete with, for example, the referential theory of tense. I agree that Iatridou's proposal does indeed carry this cost. But the argument that follows does not really need to accept the details of Iatridou's proposal. All that is necessary for the argument to follow is that there is some story according to which the presupposition of the subjunctive is attributable to the presupposition of some piece of morphology present in the subjunctive conditional, and not to the semantics of *if* itself.

context to make salient either the epistemic or the metaphysical modal parameters. I suggested that Kratzer's view leaves unexplained why the epistemic reading is salient in the contexts in which the antecedent is compatible with the common ground, and why the metaphysical reading is salient when the antecedent is incompatible with the common ground. In the Iatridou case, we are asking the context to make salient either the pastness or the counterfactuality. This is much easier to do: for example, (72) is trivially false if *left* is read as a real past tense.

This proposal allows Iatridou to claim that '... the meaning of a CF conditional is exactly the meaning of a non-CF conditional augmented by (57)', where (57) is the Exclusion Feature -- i.e. a presupposition-triggering feature that shows up as past tense morphology in subjunctives. (Iatridou 2000, p. 247).

I think Iatridou's argument is conclusive. The proposal explains why conditional sentences with subjunctive morphology carry the presuppositions that they do. Although it does not, immediately, explain why conditional sentences with indicative morphology carry the presuppositions that they do, there are ways to fill in that part of the story. Perhaps the simplest is to appeal to Heim's principle of maximizing presuppositions: then, the utterance of the indicative conveys that the presuppositions of the subjunctive are *not* satisfied, which gives the desired presupposition of the indicative.⁸⁷

It appears, at first glance, that in Iatridou's proposal we have everything we wanted. The

87. There are other ways of extending Iatridou's account, besides the appeal to maximizing presuppositions. See, for example, Schlenker 2005. It suffices to posit a competition mechanism between indicative and subjunctive morphology. So we can say that the subjunctive morphology is marked, and the indicative unmarked (cf. Schlenker 2005). Then, the absence of a subjunctive morphology conveys the presupposition that subjunctive morphology is not warranted, and therefore, that the antecedent is not outside the common ground. (see Sauerland 2003 and Schlenker 2005 for more on the topic of antipresuppositions).

posit of indicative and subjunctive features (whatever the details) allows for a division of labor: the semantics of the conditional does its part, and the tense and aspect features responsible for the distinctive morphology do theirs, viz. of explaining why indicatives and subjunctives carry the presuppositions that they do. Iatridou provides a satisfying answer to the presupposition question. But note, crucially, that Iatridou's story provides a satisfying explanation only when coupled with an appropriately unified semantics of the conditional. Iatridou herself says: '... the meaning of a CF conditional is exactly the meaning of a non-CF conditional augmented by (57) (*viz. the exclusion feature*)'. But Iatridou's proposal does not show, by itself, that this is the case, because not just any unified theory of the conditional will provide the desired explanation of presuppositions. Let me explain.

Consider the combination of the Iatridou's feature mechanism with Kratzer's semantics. The problem is this: suppose the antecedent is compatible with the common ground, but, in our context, we are interested in the metaphysical modality. Then, it seems, we can utter 'if P, then Q', with *indicative* morphology, but with the *metaphysical* modal base and ordering source *f* and *g*. (again, the riverboat scenario is a good example here). Such readings are not attested. The problem, again, is in placing the choice between metaphysical and epistemic readings in the hands of the context of the conversation.

The feature story explains why sentences in the indicative mood presuppose that the antecedent is live, and why sentences in the subjunctive mood presuppose that the antecedent is not live, but it does *not* explain why indicatives cannot carry *subjunctive meaning*, where this is understood as some semantics that appeals to the metaphysical modality, as distinct from the

epistemic modality in play with indicatives. Appeal to presupposition-triggering features only solves the presupposition problem when combined with a semantics of the conditional that really gives the conditional *the same meaning* across all contexts.

In particular, Iatridou does not show 'that it is possible to attribute the difference in meaning to a systematic difference in verbal morphology.' (Iatridou 2000, p. 232) -- that can only be done by producing an appropriate meaning which, combined with Iatridou's theory, will account for the difference in meaning between indicatives and subjunctives. *So, for example, Kratzer's view, combined with Iatridou's story, attributes the difference in meaning to two factors: verbal morphology, and context (which makes epistemic or metaphysical reading salient).*

A semantics that gives *if* unified, non-context-sensitive truth-conditions, when combined with Iatridou's story, would provide a complete solution. Why? because there would only be one conditional meaning. So let us look at Stalnaker-inspired approaches.

6 My view

In section 5 we saw that while the feature-driven presupposition mechanism explains why conditionals with indicative and subjunctive morphology carry the presuppositions that they do, this is not enough: problems persist if the underlying semantics of the conditional is too disjunctive (we concentrated on one kind of disjunctiveness: Kratzer's contextualist view, with (74) below, we shall see another kind).

There is a view that *is* unified enough to solve the presupposition problem. Although, as

we shall see, it is unacceptable on other grounds, it is a good starting point.⁸⁸ Let's call it the *simple view*. The view is as follows:

(73) $\llbracket \text{if } P, \text{ then } Q \rrbracket^w = 1$ iff the closest P -world to w is a Q -world.

(here and throughout, I only consider the simple case where Q is not context-sensitive -- for the more complex case, see the treatment in the whether-conditionals paper.)

That is, the semantics of the conditional is *the same* as the semantics of the Lewis-Stalnaker counterfactual: the selection function searches the entire universe of worlds for a closest antecedent-world. Combined with the feature-driven explanation of conditional presuppositions, the view delivers the desired predictions: when the closest P -world is in the common ground, the conditional has indicative morphology, when the closest P -world is not in the common ground, the morphology is subjunctive. The question why subjunctive meaning does not appear with indicative morphology does not arise, because the conditional has just one kind of meaning -- that given in (73). Success is due to the fact that the semantics of *if* is not disjunctive (in particular, not context-sensitive).

Unfortunately, the simple view must be mistaken. One way to see the problem is through Gibbard stand-offs. Recall, briefly, the Sly Pete story:

Sly Pete. Sly Pete and Mr. Stone are playing poker on a Mississippi riverboat. It is now up to Pete to call or fold. My henchman Zack sees Stone's hand, which is quite good, and signals its content to Pete. My henchman Jack sees both hands, and sees that Pete's hand is rather low, so that Stone's is the winning hand. At this point, the room is

88. This view is rather similar to the view laid out in Stalnaker 1975. But the 1975 includes the claim that the selection function is context-sensitive. I'll comment on Stalnaker's 1975 view below.

cleared. [...] Zack knows that Pete knew Stone's hand. He can thus appropriately assert "If Pete called, he won." Jack knows that Pete held the losing hand, and thus can appropriately assert "If Pete called, he lost." (Gibbard (1981), p. 231).

Naturally, Pete folds. Now, concentrate on Zack's conditional. It is true, we may suppose, despite the fact that the nearest world in which Pete calls, he loses. And, when the news that Pete folded becomes available, Zack may truthfully say: had Pete called, he would have lost. If (73) were the right semantics, Zack's original indicative and the later subjunctive would be inconsistent. But they are not, so the simple view is out.

One way of learning the lesson of the riverboat scenario, if one wants to keep close to the simple view, is to posit that the indicative is *contextually restricted*: the indicative does not search for the closest-antecedent world in the whole universe of worlds, but only in the set of worlds consistent with what is taken for granted -- in the common ground. Some philosophers sympathetic to Stalnaker's early views have made this adjustment -- e.g. Heim 1992, Nolan 2003. The resulting view would be:

(74) $\llbracket \text{if } P, \text{ then } Q \rrbracket^{C,w} = 1$ iff the closest P -world to w in C , the common ground, is a Q -world.

(74) works well as a semantics of the indicative.⁸⁹ But what if we take (74) to be a unified semantics? Now a real issue arises: the semantics in (74) cannot account for counterfactual readings. To make the modal base C context-dependent seems to go straight back to Kratzer's so-

89. Another solution is what Stalnaker proposes in 1975: to make the selection function, i.e. the similarity relation, context-sensitive. One may then posit that worlds in the common ground are always closer than the worlds outside it. So far as I can see, this view is a formal variant of (74): it 'folds' the explicit contextual restriction into the similarity relation. But it is philosophically less satisfying: after all, our intuitions about similarity, whether by appeal to Lewis' world-ranking procedure, or by appeal to causality (Kment), elicit a context-independent relation. So it is plausible that we should separate the purely metaphysical relation of similarity from epistemic contextual effects, as done in (74).

lution, which has already been rejected. So this proposal, too, must be rejected. What to do?

The first thought is to distinguish two kinds of context-dependence. One model is illustrated by Kratzer's view: here the context supplies the value of a parameter of evaluation -- in some cases the epistemic, in some cases the metaphysical ordering source and modal base. Another model is illustrated by (74): here the context supplies values of a single parameter: the epistemic context. The second kind of context-dependence is no threat. So can one propose a unified semantics along the lines of (74)?

Here is a natural thought: the conditional aims to find a closest antecedent-world in *the minimal context compatible with the antecedent*. What I am suggesting is similar to presupposition accommodation: the audience accommodates the presupposition by adjusting the context minimally to make the presupposition true. But what I am suggesting is *not* presupposition accommodation: the accommodation of a counterfactual antecedent does not survive even if the conditional is accepted (accepting 'had Oswald not killed Kennedy, someone else would have' does not result in a context in which it is taken for granted that Oswald did not kill Kennedy). Perhaps there is a distinct mechanism at play here, but, without knowing much about its nature, I propose instead to enrich the semantics of the conditional so that it delivers an equivalent result. Here is how to do it.

The evaluation procedure goes as follows: first, a minimal context compatible with the antecedent is found, and then the semantics searches for a closest antecedent-world in that minimally adjusted context. In the case of the indicative, the initial context -- the common ground -- need not be revised, and so the first step idles. But in the subjunctive case, the context

needs to be expanded, and so the first step does some work. Slightly more formally, the semantics can look as follows:

(75) $\llbracket \text{if } P, \text{ then } Q \rrbracket^{C, w} = 1$ iff the closest P -world to w in $C+P$ is a Q -world.

where $C+P$ is a minimal revision of common ground C that makes C compatible with P .

(so '+' is the revision function)

In at least a minimal sense, (75) provides an answer to the presupposition question. But the proposal in (75) is essentially more complicated than (74), since it introduces a new notion of minimal revision of context.⁹⁰ So I think it is fair to say that one ought to view (75) with suspicion, and the plausibility of (75) remains to be proven. In the remainder of this paper, I will make one important step toward that goal.

7 Context-dependence of subjunctives

Surely the main issue with (75) is that it may be justly suspected of being disjunctive beneath the surface, and thus not achieve the promised explanatory gain. So, if $C+P=C$ when C is compatible with P , and $C+P=\emptyset$ when C is incompatible with P , one might think that the notion of minimality is not doing any work: for any C , the function $C+P$ has only two possible values, and to say that the function expresses minimal revision is just to say that one of the values (C) is ranked below the other (\emptyset). In this case, the advantage of (75) over Kratzer is min-

90. What is minimal revision? I am, for now, happy with the following picture: if $C \cap \llbracket P \rrbracket = \emptyset$, $C+P = C \setminus P$, i.e. the set of propositions in C less the proposition expressed by P . So the set of worlds *expands* when P is incompatible with C .

imal: Kratzer offers contextual variation between metaphysical and epistemic readings, and runs into problems when explaining why metaphysical readings are not available when the antecedent is compatible with the common ground. But the same charge of arbitrariness can be leveled against (75): it just stipulates that C+P turns to the metaphysical scope when P is not compatible with C. Can this charge of arbitrariness be answered? I think it can, in a surprising way: (75) is not *ad hoc* because subjunctives are context-sensitive in such a way that C+P has more values than just C and \emptyset , and so minimality is a contentful notion. Let me explain.

First, it is important to distinguish different ways of being context sensitive. There is a traditional line of thought, sanctioned by Lewis, that the similarity metric of counterfactuals is sensitive to the goals of a conversation. The standard example is Quine's:

(76) If Caesar were in command in Korea, he would use the atom bomb.

(77) If Caesar were in command in Korea, he would use catapults.

(Quine 1960, 221), also in Lewis 73, p. 66-7

Here, the idea is that somehow (76) might be true in some contexts, and (77) in others -- depending on the interest of the participants of the conversation. So this is a familiar kind of context-sensitivity. But the kind of context-sensitivity I am attributing to subjunctives is of a fundamentally different kind.

The Lewis-context-sensitivity is a sensitivity to the goals of the conversation. Participants in a conversation, sharing the same information, can, depending on their interests, steer the conversational context to make it the case that one or another similarity relation is salient. I am suggesting something very different. The kind of context-sensitivity I am proposing is not

a matter of salience. According to (75), the conditional is sensitive to the epistemic context of the conversation: to what is known, or taken for granted. So (75) is insensitive to interest and conversational salience. The participants in a conversation sharing the same information have, according to (75), a fixed interpretation for their conditionals, no matter what their conversational goals.

The hallmark of context-sensitivity of indicatives are the Gibbard stand-off pairs that we saw already. The riverboat scenario shows how an indicative is dependent on epistemic context: depending on what you know, a given indicative can be true for you or false for you, even when all the relevant portions of the world remain the same.⁹¹ The riverboat scenario dramatizes this difference: Zack and Jack utter apparently contradictory indicatives, yet both are true, because Zack and Jack know different things about the poker game. I want to claim that subjunctives display essentially the same kind of context-sensitivity: *they are also dependent on what we know*. This is a radical view; what's the evidence?

Swanson (forthcoming) discusses an example of subjunctive stand-offs:

Suppose that Al, Bert, Carl, Dawn, Eve, and Fran are siblings. It's common ground that their parents are considering taking a trip to London, and that if they go they will bring Al, Bert, Carl, and exactly one of Dawn, Eve, and Fran. From different vantage points, Al and Bert witness a conversation between their parents and at least some of their siblings. From his vantage point, Al sees Dawn, Eve, and Fran walk into the room, sees Fran leave, and hears another sibling leave. He then hears their parents telling either

91. So, unlike Gibbard himself, who offered the stand-offs as an argument for a non-truth-conditional semantics of the indicative, I think that the lesson of the riverboat scenario is that indicatives are information-sensitive.

Dawn or Eve (he's not sure which) that they will take her if they go. From his vantage point, Bert sees Dawn, Eve, and Fran walk into the room, hears a sibling leave, and sees Eve leave. He then hears their parents telling either Dawn or Fran (he's not sure which) that they will take her if they go. Later, Al says to Carl:

(23) I want to go to London. We would see Big Ben, and the Tate Modern. And if Dawn weren't with us, Eve would be, although Fran wouldn't be.

And Bert says to Carl:

(24) I want to go to London. We would see Big Ben, and the Tate Modern. And if Dawn weren't with us, Fran would be, although Eve wouldn't be.

Al and Bert's subjunctive conditionals constitute a Gibbardian stand-off. (Swanson (forthcoming), p. 7)

One problem with the example is that the future less vivids are not felicitous: in the context as described, Al ought to say "And if Dawn is not with us, ...", and Bert ought to say "And if Dawn is not with us...". But there is a more basic problem as well. Suppose Al and Bert go to London, with Dawn. Now transpose Al's and Bert's utterance into past tense. Now Al says:

(78) If Dawn were not with us, Eve would be.

and Bert says:

(79) If Dawn were not with us, Fran would be.

Do we have any inclination to say that both Al's and Bert's utterances are true? Not at all. So Swanson's example does not work.

Here is a better *counterfactual* Gibbard stand-off:

The scale

The experimenter is about to conduct, in succession, two experiments. First, he will put two weights, A and B, onto the left and the right cup of a simple pharmacy scale (the kind where two cups are suspended from a rod that is itself affixed in the middle to a standing arm). In the second experiment, to be carried out immediately after the first, the experimenter will switch the weights: he will put weight A into the right cup, and weight B into the left.

Now, we have two observers, X and Y. Crucially, X knows that weights A and B have the same mass (and so weight), but does not know whether the arms of the scales are of equal length. Y, on the other hand, knows that the arms are exactly equal, but does not know whether A and B have the same mass. (Let's further assume that enough is known about the set-up to make it the case that just these two parameters -- length of arms and the weight of weights determine whether the scale will tip one way or the other -- so, no atmospheric effects, no hidden magnets, it is known that the cups of the scale are of equal weight, etc. -- as a result, as a matter of nomic necessity, the scales will remain perfectly balanced in both experiments).

Now, before the experimenter embarks on his manipulations, X says:

(80) If the right cup goes down in the first experiment, then the *right* cup will also go down in the second experiment.

Naturally, if the right cup goes down, or rather if the right cup were to go down, X would conclude that the right arm is shorter than the left arm.

Likewise, before the experiments, Y says:

- (81) If the right cup goes down in the first experiment, then the *left* cup will go down in the second experiment.

Naturally, if the right cup goes down, or rather if the right cup were to go down, Y would conclude that weight A is lighter than weight B. So far, we have an *indicative* Gibbard stand-off.

Now the first of the experiments is performed, but the experimenter *only tells X and Y that the right cup did not go down*.⁹² X can then say, truly:

- (82) Had the right cup gone down in the first experiment, then the *right* cup would also have gone down in the second experiment.

and Y can say, truly:

- (83) Had the right cup gone down in the first experiment, then the *left* cup would have gone down in the second experiment.

Hence, subjunctives can produce Gibbard stand-offs, and so the semantics in (75) is to this extent justified: the minimal change of the epistemic context C that (75) posits is non-trivial.

In retrospect, the effect of Gibbard's indicative stand-offs -- to convince the philosophical community that indicatives are context-sensitive while subjunctives are not -- rests on a lucky choice of example. What is distinctive about the riverboat scenario is that on Jack's side there is a physical law -- to the effect that if the cards in both hands are as Jack saw them, they would not suddenly change. For this reason Jack's indicative survives as a subjunctive -- or, as

92. This complication is needed to avoid the difficult question of what X and Y would say if they saw the result of the first experiment. Once one sees that the scale has remained balanced, it is not clear what to say about 'Had the right cup gone down,...' -- this would be a counter-legal. There is nothing wrong with counter-legals, but I want to keep these issues separate, hence the complication that limits X's and Y's knowledge of the outcome of the experiments.

one might put it, underwrites a later subjunctive. On the other hand, Zack's indicative is not underwritten by a physical law, but only by an empirical generalization -- to the effect that Sly Pete always acts in his best interest, or something like that. For this reason, we have a strong intuition that Zack's indicative does not survive as a subjunctive once the news comes in that Sly Pete folded. But this situation is rather special. The central insight behind my scale example is that there is nothing in the original Gibbard scenario that demands that one of the observers should rely on a physical law, and the other on something weaker. The scale scenario is a scenario in which both observers rely on laws of equal strength, so to speak, and thus both survive as subjunctives once the news comes in that the antecedent is false.

Thus, the general pattern of a subjunctive stand-off is as follows: observer X knows of a law N such that 'If P, then Q' is underwritten by N. Observer Y knows of a law M such that 'if P, then \sim Q' is underwritten by law M. But since both laws obtain, P is nomically impossible. If X and Y know enough, but not too much, a subjunctive stand-off results.

8 Conclusion

The argument of sections 2-6 points to a theory like (75):

$$(75) \quad \llbracket \text{if } P, \text{ then } Q \rrbracket^{C,w} = 1 \text{ iff the closest } P\text{-world to } w \text{ in } C+P \text{ is a } Q\text{-world.}$$

where $C+P$ is a minimal revision of common ground C that makes C compatible with P .

The theory is a unified theory of the conditional. The semantics of the conditional is presup-

position-less in the sense that, according to it, conditional LFs are everywhere defined.⁹³ In order to derive the presuppositions of indicatives and subjunctives, we need a revision of labor: a theory like Iatridou's, that associates indicative and subjunctive presuppositions with the relevant morphology found in indicative and subjunctive sentences. (75), indeed, completes Iatridou's explanation of conditional presuppositions: it supplies *a single meaning* such that the difference between indicatives and subjunctives can be accounted for by appeal to the differences in morphology.

But the context revision function at work in (75) is idle unless it can be shown that subjunctives require non-trivial revision of context. The argument in section 7 shows that this is indeed the case.

The unique feature of (75) is that both indicatives and subjunctives are sensitive to the epistemic context. In this sense indicatives and subjunctives both report on the same modal realm.

93. There is a possibility that there are cases in which the minimal revision function fails to return a value. If such cases exist, (75) will carry a semantic presupposition (to the effect that such cases do not obtain.)

BIBLIOGRAPHY

- Adams, Ernest. 1970. Subjunctive and indicative conditionals. *Foundations of Language* 6(1). 89-94.
- Adams, Ernest (1975). *The Logic of Conditionals*. Dordrecht: Reidel
- Ahmed, Arif & Price, Huw (forthcoming). Arntzenius on 'Why ain't cha rich?'. *Erkenntnis*:-.
- Aloni, Maria and Roelofsen, Floris (2011), Indefinites in Comparatives. *Proceedings of SALT* 21: 19-38, 2011
- Barker, Stephen (1997). Material implication and general indicative conditionals. *Philosophical Quarterly* 47 (187):195-211.
- Barnett, David (forthcoming). Future Conditionals and DeRose's Thesis. *Mind*:-
- Beck, Sigrid. (2010). Quantifiers in than-clauses. *Semantics and Pragmatics* 3. 1-72.
- Bennett, Jonathan (2003). *A Philosophical Guide to Conditionals*. Oxford: Clarendon Press
- Cantwell, John (2006). The logic of dominance reasoning. *Journal of Philosophical Logic* 35 (1).
- Cantwell, John (2008). Indicative conditionals: Factual or epistemic? *Studia Logica* 88 (1).
- Cantwell, John (2010). On an alleged counter-example to causal decision theory. *Synthese* 173 (2).
- Chalmers, David J. (2006). Two-dimensional semantics. In E. Lepore & B. Smith (eds.), *Oxford Handbook of the Philosophy of Language*. Oxford University Press.
- Edgington, Dorothy (1995). On conditionals. *Mind* 104 (414):235-329.
- Edgington, Dorothy (2011). Conditionals, causation, and decision. *Analytic Philosophy* 52 (2):75-87.

- Eells, Ellery and Skyrms, Brian, (eds) (1994). *Probability and Conditionals: Belief revision and rational decision*. Cambridge: Cambridge University Press.
- Egan, Andy. 2007. Some Counterexamples to Causal Decision Theory. *Philosophical Review*, 116: 93–114.
- von Fintel, Kai (1998), The presupposition of subjunctive conditionals. In Uli Sauerland & Orin Percus (eds.), *The interpretive tract*. (MIT Working Papers in Linguistics 25), 29–44. Cambridge, MA: MITWPL.
- Gibbard, Allan (1981a): Two recent theories of Conditionals. In Harper, Stalnaker, and Pearce 1981, pp. 211–47.
- Gibbard, Allan and Harper, William (1981b). Counterfactuals and Two Kinds of Expected Utility. In William Harper, Robert Stalnaker, and Glenn Pearce, eds., *Ifs: Conditionals, Belief, Decision, Chance, and Time*, pp. 153–190. Dordrecht: Reidel.
- Gillies, Anthony S. (2007). Counterfactual scorekeeping. *Linguistics and Philosophy* 30 (3):329–360.
- Gillies, Anthony S. (2009). On truth-conditions for if (but not quite only if). *Philosophical Review* 118 (3).
- Greco, Daniel (forthcoming). Probability and prodigality. *Oxford Studies in Epistemology*, Volume 4.
- Harper, W. L., Stalnaker, R., and Pearce, C. T. (eds.) 1981: *Ifs*. Dordrecht: Reidel.
- Heim, Irene (1992). Presupposition Projection and the Semantics of Attitude Verbs. *Journal of Semantics* 9 (3):183–221.

- Heim, Irene. (2006). 'Remarks on comparative clauses as generalized quantifiers.' Ms. MIT
- Hiddleston, Eric (2005). A causal theory of counterfactuals. *Noûs* 39 (4):632-657.
- Higginbotham, James (1986). Linguistic theory and Davidson's program in semantics. In E. Lepore (Ed.), *Truth and Interpretation: Perspectives on the philosophy of Donald Davidson* (pp. 29-48). Oxford: Basil Blackwell.
- Iatridou, Sabine. (2000). Grammatical ingredients of counterfactuality. *Linguistic Inquiry* 31 (2): 231-270.
- Jackson, Frank (1987). *Conditionals*. Oxford: Blackwell
- Joyce, James M. (1999). *Foundations of causal decision theory*. Cambridge, UK: Cambridge University Press.
- Joyce, James M. (2007). Are newcomb problems really decisions? *Synthese* 156 (3):537 - 562.
- Kaufmann, Stefan (2004). Conditioning against the grain. *Journal of Philosophical Logic* 33 (6).
- Kennedy, Christopher (2007). Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy* 30 (1):1 - 45.
- Klinedinst, Nathan & Daniel Rothschild (forthcoming). Connectives without truth-tables. *Natural Language Semantics*.
- Kment, Boris (2006). Counterfactuals and the analysis of necessity. *Philosophical Perspectives* 20 (1):237-302.
- Kment, Boris (2010). Causation: Determination and difference-making. *Noûs* 44 (1):80-111.
- Kratzer, Angelika (1986). Conditionals. *CLS* 22
- Kratzer, Angelika (2012). *Modals and Conditionals*. Oxford Studies in Theoretical Linguistics

36. Oxford, UK: Oxford University Press.

Lasnik, Peter (1996). Adnominal conditionals. In T. Galloway & J. Spence (eds.), *Papers from Semantics and Linguistic Theory VI*. CLC Publications.

Lewis, David (1976). Probabilities of Conditionals and Conditional Probabilities.'

Philosophical Review, 85, pp. 297-315.

Lewis, David (1981)a. 'Why ain't cha rich?' *Noûs* 15 (3):377-380.

Lewis, David (1981)b. Causal decision theory. *Australasian Journal of Philosophy* 59 (1):5–30.

McGee, Vann (1985). A counterexample to modus ponens. *Journal of Philosophy* 82 (9):462-471.

McGee, Vann (1989). Conditional probabilities and compounds of conditionals.

Philosophical Review 98 (4):485-541.

Nolan, Daniel (2003). Defending a possible-worlds account of indicative conditionals. *Philosophical Studies* 116 (3).

Pollock, John (forthcoming). A recourse-bounded agent addresses the Newcomb problem.

Synthese.

van Rooij, Robert, Comparatives and quantifiers. In *Empirical Issues in Syntax and Semantics* 7,

O. Bonami & P. Cabredo Hofherr (eds.), 2008, pp. 423-444.

Schlenker, Philippe (2004). Conditionals as Definite Descriptions (A Referential Analysis), *Research on Language and Computation*, 2, 3: 417-162, 2004

Schlenker, Philippe (2005). The Lazy Frenchman's Approach to the Subjunctive (Speculations on Reference to Worlds and Semantic Defaults in the Analysis of Mood). *Proceed-*

ings of Going Romance XVII.

Stalnaker, Robert (1975). Indicative conditionals. *Philosophia* 5 (3).

Stalnaker, Robert (1984). *Inquiry*. Cambridge, Mass.: The MIT Press.

Schwarzschild, Roger (2008). The semantics of comparatives and other degree

constructions. *Language and Linguistics Compass*, 2008

Schwarzschild, R. & Wilkinson, K. (2002). Quantifiers in comparatives: A semantics of degree

based on intervals. *Natural Language Semantics*, 10:1-41.

Swanson, Eric (forthcoming). Subjunctive biscuit and stand-off conditionals. *Philosophical Studies*-.

Wasserman, Ryan (2006). The future similarity objection revisited. *Synthese* 150 (1):57 - 67.

Wedgwood, Ralph (2011). Gandalf's solution to the Newcomb problem. *Synthese* (Online First), March 14 :1-33.

Williamson, Timothy (2006). Indicative versus subjunctive conditionals, congruential versus non-hyperintensional contexts. *Philosophical Issues* 16 (1):310 333.